

THE TECHNICAL COOPERATION PROGRAM

Subcommittee on non-atomic military research and development

Group HUM - Human Resources and Performance

Human Performance Modeling in Military Simulation: Current State of the Art and the Way Ahead

Report of Action Group 19 – Representation of Human Behavior

26 July 2004

TR-TTCP/HUM/02/02

This document contains information authorized under the auspices of The Technical Cooperation Program (TTCP) for unlimited release and distribution..

Report Documentation Page			<i>Form Approved OMB No. 0704-0188</i>	
<p>Public reporting burden for the collection of information is estimated to average 1 hour per response, including the time for reviewing instructions, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing the collection of information. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing this burden, to Washington Headquarters Services, Directorate for Information Operations and Reports, 1215 Jefferson Davis Highway, Suite 1204, Arlington VA 22202-4302. Respondents should be aware that notwithstanding any other provision of law, no person shall be subject to a penalty for failing to comply with a collection of information if it does not display a currently valid OMB control number.</p>				
1. REPORT DATE 03 SEP 2003	2. REPORT TYPE N/A	3. DATES COVERED -		
4. TITLE AND SUBTITLE 2002 TTCP HUM Group Meeting: Human Performance Modeling in Military Simulation AG-19: Current State of the Art and the Way Ahead				
6. AUTHOR(S) 				
5a. CONTRACT NUMBER 5b. GRANT NUMBER 5c. PROGRAM ELEMENT NUMBER 5d. PROJECT NUMBER 5e. TASK NUMBER 5f. WORK UNIT NUMBER				
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) TTCP				
8. PERFORMING ORGANIZATION REPORT NUMBER				
9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)				
10. SPONSOR/MONITOR'S ACRONYM(S)				
11. SPONSOR/MONITOR'S REPORT NUMBER(S)				
12. DISTRIBUTION/AVAILABILITY STATEMENT Approved for public release, distribution unlimited				
13. SUPPLEMENTARY NOTES See also ADM201725, DOC-TTCP/HUM/02/09, TTCP HUM Group Annual Meeting Minutes Jul 15-19, 2002., The original document contains color images.				
14. ABSTRACT				
15. SUBJECT TERMS				
16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT UU	18. NUMBER OF PAGES 143
a. REPORT unclassified	b. ABSTRACT unclassified	c. THIS PAGE unclassified	19a. NAME OF RESPONSIBLE PERSON	

The Technical Cooperation Program HUM Action Group 19 Members

Australia

Simon Goss
Defence Science and Technology
Organization
Aeronautical and Maritime Research
Laboratory
506 Lorimer Street, Fishermans Bend
Victoria
PO Box 4331, Melbourne VIC 3001

Canada

Brad Cain
Defence Research and Development Canada
Defence and Civil Institute of
Environmental Medicine,
1133 Sheppard Avenue West,
Toronto, ON M3M3B9

Len Kerzner
Directorate of Manpower Analysis
Chief Personnel Policy Planning and
Resource Management
NDHQ, 101 Colonel By Dr.
Ottawa, ON

United States

Harold Hawkins, Chairman AG-19
Cognitive and Neural Science &
Technology Division
Office of Naval Research
800 N. Quincy Street
Arlington, VA 22217

Ruth P. Willis
US Naval Research Laboratory
Advanced Information Technology Branch
4555 Overlook Ave., SW
Washington, DC 20375-5337

Michael J. Young
Air Force Research Laboratory
Human Effectiveness Directorate
2698 G Street
Wright-Patterson AFB, OH 45433-7604

Denise Lyons
Naval Air Warfare Center
Training Systems Division
12350 Research Parkway
Orlando, FL 32826-3224

United Kingdom

Colin Sheppard
Defence Science and Technology
Laboratory
DSTL Analysis Maritime Studies
DSTL Portsdown West
Portsdown Hill Road
Fareham PO17 6AD, UK

Executive Summary

The use of modeling and simulation (M&S) is pervasive throughout military establishments. M&S tools are used to support the development and acquisition of new systems, to evaluate existing operational plans, to develop new war-fighting concepts, and to train tactics. This report examines the requirements for human performance modeling within the military, assesses the state of the practice in current operational models, documents ongoing human performance research and development (R&D) projects, identifies shortfalls in the competence of available models, and recommends a roadmap of research to eliminate these shortfalls.

Chapter 1 examines human performance modeling requirements. In this chapter we provide examples of operational or training shortfalls and needs arising from the military user and technical communities that may be overcome by adding high fidelity Human Performance modeling to existing tools. Wherever possible, formal needs have been referenced. The chapter examines human performance modeling requirements within the domains of simulation-based acquisition, training, mission rehearsal, and operations research. A major conclusion to be drawn from this chapter is that explicit requirements for human performance models are uncommon in the TTCP military M&S establishments. Where explicit requirements do exist, metrics are rarely provided that would enable one to determine whether the requirements have been satisfied.

Chapter 2 focuses on human performance modeling capabilities currently used in simulations for military training and analysis. The chapter is divided into two sections. The first examines the use of human modeling within computer generated forces, such as ModSAF, Soar, JSAT, and the Close Combat Tactical Trainer. The second section examines the use of intelligent agents as representations of military operators. The most striking conclusion that comes from this review is the limited number of human performance modeling applications currently in use in the TTCP military communities.

Chapter 3 documents sponsored research aimed at enhancing the competence, realism, and military application potential of human performance models. These summaries extend from basic science to advanced research and development activities and represent on-going research from the United States, United Kingdom, Canada, and Australia during the period of 2000-2002.

UNCLASSIFIED

The conclusion to be drawn from this analysis is that while considerable work is underway in human performance modeling, the field remains relatively immature.

Chapter 4 focuses on the disparity between identified needs and existing capabilities. There is a broad consensus among military simulation organizations that existing military simulations do not possess sufficient fidelity, particularly in the area of human performance, to answer the questions the military must answer in order to be fully prepared for current and future operations. The shortfalls in human performance modeling were assigned to two broad categories: the behavioral competence and realism of models, and the technologies supporting their application. A critical agenda item for the TTCP M&S community in the years ahead is to study techniques to enhance the behavioral and cognitive competence of human performance models along dimensions critical to military operations. AG-19 recommends that TTCP follow-on activity be captured and managed under the TP-2 (Training) umbrella, starting with a collaborative activity involving the US and CA, although it is hoped that UK, NZ and AU would also participate.

Table of Contents

<u>Chapter 1: Human Performance Modeling Requirements</u>	1
<u>Simulation-Based Acquisition</u>	4
<u>Training</u>	8
<u>Mission Rehearsal</u>	14
<u>Operations Research</u>	18
<u>Conclusion</u>	19
<u>References</u>	20
<u>Chapter 2. Assessment of Current Operational Capabilities in Human Performance Modeling</u>	22
<u>Current Capabilities for Computer Generated Forces</u>	22
<u>ModSAF</u>	22
<u>Soar</u>	23
<u>JSAF</u>	25
<u>Close Combat Tactical Trainer</u>	26
<u>Australian Experiences with Agent Oriented Simulation</u>	26
<u>Agent Description</u>	26
<u>Representation of Team Structures</u>	29
<u>Agent Oriented Design for Simulators</u>	31
<u>REFERENCES</u>	32
<u>Chapter 3 Human Performance Research and Development Projects</u>	35
<u>United States</u>	35
<u>United States Army</u>	35

UNCLASSIFIED

<u>Human Behavioral Modeling Program</u>	35
<u>Related Publications</u>	36
<u>SOAR Technology</u>	36
<u>Natural Selection Inc.</u>	37
<u>Psychometrix Associates, Inc.</u>	39
<u>Related Publications</u>	39
<u>Micro Analysis and Design</u>	40
<u>Related Publications</u>	41
<u>Institute for Creative Technologies</u>	41
<u>Related Publications</u>	42
<u>The United States Navy and Marine Corps</u>	42
<u>Basic (6.1) Research</u>	43
<u>Robustness in Behavioral Modeling</u>	43
<u>Extending COGNET CGF Human Modeling Capabilities</u>	44
<u>A Hybrid Architecture for Human Performance Modeling</u>	47
<u>Extending ACT-RPM to Multitasking Environment</u>	48
<u>Graphical Visualization of Situational Awareness for Computer Generated Forces</u>	49
<u>Testing an interface to explain cognitive models to experts</u>	51
<u>Developing Computational Models of Naturalistic Decision Making</u>	51
<u>Research into the Mathematical Modeling of Human Behavior</u>	52
<u>Research on Autonomous Synthetic Entities</u>	53
<u>Applied (6.2) Research</u>	55
<u>Team Training with Synthetic Teammates (SYNTHERs)</u>	55
<u>Study of Behaviors for Synthetic Teammates in Team Training</u>	55
<u>Synthetic Force Applications of Cognitive Modeling Technology in Navy Training Systems</u>	56

UNCLASSIFIED

<u>Composable Behaviors in JSAF</u>	57
<u>Cognitive and Analytic Modeling Approaches for Tutoring Systems (CAATS)</u>	58
<u>Advanced (6.3) Research and Development</u>	59
<u>Intelligent Agents to Support Real-Time Exercise Control and Data Collection in Distributed Training Exercises</u>	59
<u>Synthetic Cognition for Operational Team Training (SCOTT)</u>	60
<u>Combat Automation Requirements Testbed (CART) Program</u>	62
<u>Agent-based Modeling and Behavior Representation (AMBR) Role Player Intelligent Controller Node (RPICN)</u>	65
<u>Agent-based Modeling and Behavior Representation (AMBR) Model Comparison</u>	66
<u>Intelligent Mission Controller Node (IMCN)</u>	68
<u>Graphical Agent Development Environment II (GRADE II)</u>	70
<u>Enhancing the Usability of Computer Generated Forces</u>	70
<u>United Kingdom</u>	71
<u>Human Factors</u>	71
<u>Capture of Realistic Behaviors for inclusion within Synthetic Environments (SE)</u>	71
<u>Supporting Effective Communication and Interpretation of Command Intent</u>	72
<u>Team Processes</u>	72
<u>Doctrine</u>	73
<u>Canada</u>	74
<u>Army Command and Control Information Systems (CCIS) Development and Training Through Simulation (12kr)</u>	74
<u>References</u>	75
<u>Tactical Aviation Mission System Simulation (TAMSS) TD (13cp)</u>	75
<u>Maritime Air Littoral Operations (MALO) Technical Demonstrator (13dj)</u>	76
<u>Cockpit Systems Integration (13ia)</u>	76

UNCLASSIFIED

<u>References</u>	77
<u>Advanced distributed mission training (ADMT) technical demonstrator (TD) (16be)</u>	77
<u>Representation of human behaviour (16bx)</u>	78
<u>Command-team skill, decision making and leadership (16ka)</u>	78
<u>Stress and coping in CF personnel during operations (16kb)</u>	81
<u>References</u>	81
<u>Human Engineering performance modelling environment (16kc)</u>	82
<u>Organizational and social issues (16kx)</u>	82
<u>STRIVE (http://www.cae.com)</u>	82
<u>References</u>	85
<u>Australia</u>	85
<u>Joint Simulation Environment</u>	85
<u>Architecture Studies</u>	85
<u>Simulation Standards Studies</u>	86
<u>Intelligent Agents Studies</u>	86
<u>Virtual Air Environment</u>	87
<u>AEW&C (Project Wedgetail) Acquisition</u>	91
<u>Maritime Surveillance</u>	92
<u>Chapter 4: The way ahead in human performance modeling</u>	97
<u>Shortfalls in the competence and realism of human performance models</u>	97
<u>High Level Cognition</u>	99
<u>Robustness</u>	101
<u>Adaptability</u>	102
<u>Behavior Moderators and Realistic Individual Differences</u>	103
<u>Modeling the Behavior of Groups, Teams, and Organizations</u>	105

UNCLASSIFIED

<u>Technologies Supporting Usability of Models and Simulations</u>	106
<u>KBS VVE&T Technology</u>	111
<u>Roadmaps of Future Developments</u>	114
<u>Knowledge Acquisition Roadmap</u>	114
<u>Validation Roadmap</u>	115
<u>Composability and Interoperability Roadmap</u>	115
<u>Overall Conclusion and Recommendation</u>	117
<u>List of Abbreviations</u>	127

Table of Figures

<u>Figure 1. Pictorial representation of modelling and simulation division by the breadth of the models involved, with higher levels being more comprehensive. The edges of the base represent the four application domains of Simulation Based Acquisition and Training (shown) as well as Warfare Analysis and Mission Rehearsal (not shown).</u>	3
<u>Figure 2. Intercept of Northern strikers.</u>	89
<u>Figure 3. First intercept of escorted Orange strikers during demonstration.</u>	90

Table of Tables

<u>Table 1. Summary of some DSTO Intelligent Agent Applications.</u>	29
--	----

Chapter 1: Human Performance Modeling Requirements

Requirements for human performance modeling vary from precise statements in formal requirements documents, such as Mission Need Statements (MNSs) and Operational Requirements Documents (ORDs), to less formal documents including assessments by technical experts in human performance and cognition and requests by users and operational commands for improved realism in exercises and training. In this chapter we provide examples of operational or training shortfalls and needs arising from the military user and technical communities that may be overcome by adding high fidelity Human Performance modeling to existing tools. Wherever possible formal needs are referenced. Unfortunately, this biases the discussion towards US programs because the US has several major modeling and simulation initiatives under way, all documented by formal requirements. It should be noted, however, that all TTCP countries have similar needs as demonstrated by the creation of M&S coordination and facilities such as the US Defense Modeling and Simulation Office (DMSO), the Synthetic Environment Co-ordination Office (SECO) in the UK, and the Canadian Forces Experimentation Centre (CFEC) as well as other government and private experimentation, simulation and research centers that deal with human behavior and performance.

The use of Modeling and Simulation (M&S) is pervasive throughout military establishments, however a review of enabling technologies compiled for the US Airforce (Andrews, et al. 1998) noted that although considerable psychological and physiological knowledge exists, little has been applied to the mainstream M&S. M&S tools are used to support the development and acquisition of new systems, to evaluate existing operational plans, to develop new war-fighting concepts, as well as to train tactics. Human performance modeling requirements vary significantly across these application domains. The complexity of human behavior continues to be a limiting factor for human behavioral research despite advances in psychology; it is likely that applied models of human behavior will continue to be engineering approximations rather than fundamental psychological models for the foreseeable future.

M&S applications vary along multiple dimensions. One dimension reflects the need to simulate different aspects of the mission space. For example, some studies are directed at

UNCLASSIFIED

evaluating the effectiveness of a specific weapon system component (e.g., the legibility of a head-up display system under a range of lighting and clutter conditions) and consequently only a few system components need to be simulated. Other studies may focus on evaluating the efficacy of a sub-system for a given mission and, accordingly, require emulation of more of the battlespace. Evaluation of a new, phase-array radar might, for example, require the emulation of an aircraft, its human pilot and opposing forces. Training applications, such as the conduct of a tank-crew scout mission, would typically require a high fidelity simulation of the weapon system, an external environment, and opposing infantry or vehicles. Finally, a more traditional Operations Research (OR) M&S application might investigate the viability of employing existing, doctrinally defined, combined Land and Air tactics, appropriate to large armor-on-armor battles in a rural environment, against a well armed, sizeable terrorist force in an urban environment. Such OR applications might require the emulation, or participation, of a large number of command and control echelons, as well as a large number of individual units.

A pyramid can be used to show two of the dimensions along which M&S applications may vary. The levels of the pyramid represent inclusiveness of the mission space, with the top of the pyramid being more inclusive, while the sides represent different application areas in which M&S tools and technologies are employed (see Figure 1). Authentic representations of human behavior are usually required to successfully accomplish study objectives, regardless of the application domain and relative inclusiveness of the simulated battlespace.



Figure 1. Pictorial representation of modelling and simulation division by the breadth of the models involved, with higher levels being more comprehensive. The edges of the base represent the four application domains of Simulation Based Acquisition and Training (shown) as well as Warfare Analysis and Mission Rehearsal (not shown).

The need for high fidelity human performance modeling has been highlighted in several recent reports. For example, the North Atlantic Treaty Organization (NATO) M&S Master Plan (2000) notes that currently available models do not accurately represent human behavior at the individual, group or team level. It further notes that new missions such as information operations necessitate the development of better representations of human behavior, particularly the decision-making process of commanders employing forces. In addition, Pew and Mavor's 1998 report from the US National Academy of Sciences found that human behavioral representation and modeling was critical for all military services due to their increasing reliance on the output of models and simulations for management, decision-making, and training activities. The study panel noted that the modeling of cognition and action of individuals and groups might be the most difficult task ever undertaken by humans and that developments in this area are still in their infancy.

In the following sections we provide examples of how human performance modeling can be used to support system acquisition, command post exercise training, studies and analysis and

mission rehearsal. Our intent is not to provide an exhaustive review of the ways that human performance models are being used within the military, but rather to provide illustrative examples of the diverse needs and requirements of those who use human performance models to accomplish military missions.

Simulation-Based Acquisition

The defense establishments of most TTCP nations have stated the need for better tools and techniques to evaluate the role of humans within future weapon systems. For example, Canada's Technology Investment Strategy for Defence Research¹ (*Technology Investment Strategy For the Next Two Decades*, 2000) notes "*the impact of the human in the loop, especially in decision-making, is difficult to model given that decisions are made in situations with many degrees of uncertainty in complex and dynamic changing environments*". This report then calls for the "*investigation of tools and capabilities for helping decision-makers choose among alternative courses of action to achieve goals*". More specifically, it states "*that models of operator workload, human performance, and function allocation must be explored and advanced to determine the most effective role for humans in future systems*". The development of such tools will lead to the creation of effective human-machine systems, enabling a responsible reduction in manning levels consistent with required levels of systems effectiveness. To support this philosophy, the Royal Canadian Air Force maintains a set of human factors engineering tools, Human Engineering Analysis Requirements Tool (HEART²), within industry that project managers call upon as required in acquisition projects; human-in-the-loop (HIL) simulation and constructive modelling for performance prediction are prominent components of this capability intended to mitigate the risk of purchasing or developing inappropriate equipment. A key component of the HEART analysis is the prediction of system performance incorporating operator capabilities and limitations. The analysis tools, performance modelling and HIL evaluations have been used in most major Canadian air acquisitions during the past 15 years.

¹ http://www.crad.dnd.ca/public/tis/tis_e.html

² http://www.crad.dnd.ca/hsr/refres_e.html

UNCLASSIFIED

Currently, modelling and simulation facilities are being applied to CF18 modernization and the definition phase of a land, command and control information system (CF C2IS). Unfortunately, the typical approach is for the analyst to specify explicitly most aspects of operator behaviour rather than having it evolve based on scenario events and sub-models of human performance based on operator state. Analysts have informally acknowledged that the unavailability of sub-models for constructive simulations is a significant obstacle to in-depth analyses of proposed systems, noting a need for rational representations of the individual characteristics listed in Table 1 and their effect on performance.

Table 1. Individual characteristics required in models of human behavioural representation.

<ul style="list-style-type: none">• fear• fatigue (mental and physical)• motivation• thermal strain and hydration• mental ability• perceptual ability• problem solving• decision making• agility	<ul style="list-style-type: none">• situation awareness• training and skill levels• mental abilities• reaction time• signal detection• vigilance• personality• psychological stress• learning and adaptation
--	--

In the TTCP HUM TP2 (Training) Strategic Review (2000), the panel notes that... “*The move to simulation-based methods as a cornerstone of material acquisition and support will require many of the enabling technologies and models that are used for training, including those that involve humans in the loop, or the simulation of humans in the loop. The ability to accurately represent the behavior of others seems essential to the achievement of valid constructive simulations. It also seems essential to realize the full potential of distributed interactive simulations for collective training since it will not always be possible to involve a large number of virtual, or live, simulations in a synthetic training exercise. Moreover,*

UNCLASSIFIED

interactive models of human behavior can possibly be used to avoid the costs, scheduling, resource commitments, and lack of control associated with training members of large teams.”

The high cost of increasingly sophisticated military equipment requires the acquisition of any new item by the Australian Defence Force (ADF) to be justified on the basis of its likely contribution to the defence of Australia in concert with other available defence assets, rather than the simple desirability of improving the capabilities of individual assets. This approach requires the comparison of quite different equipment solutions to particular defence challenges. Australian military operational analysis work in the past has usually addressed the engagement level, with correspondingly small numbers of entities interacting.

In the US, there is an official requirement to use M&S throughout the acquisition life cycle to reduce cost, schedule, and performance risk; to reduce time between concept definition and delivery of system; and to improve system performance³. Early in the acquisition process, M&S is used to assess mission area, identify cost drivers, analyze requirements and assess risk. During the engineering and manufacturing phase of acquisition, M&S is used to perform cost-benefit trade-off among alternatives as well as in engineering studies of system or component performance. During production and fielding, M&S is used to evaluate engineering design changes both of the system and the manufacturing process. The Canadian Forces Army Simulation Centre (ASC) provides the CF Land Force with advice based on war-game analyses. The military client has not indicated a requirement to incorporate HBR into their simulations. While ASC acknowledged HBR models would be nice to have on hand, they felt that the levels of fidelity and validation of such models as were available were insufficient for practical applications. ASC indicated that there were a large number of assumptions and limitations within the existing simulation environments such that it requires significant effort to get client approval or acceptance of an analysis tool without increasing the effort by adding the uncertainties of HBR models. ASC staff acknowledged that while they would not now consider implementing new HBR for their primary analysis, HBR constraints would make a useful addendum for performing sensitivity analyses on their primary investigation conclusions.

³ DoD 5000.2-R, Mandatory Procedures for Major Defense Acquisition Programs (MDAPS) and Major Automated Information Systems (MAIS) Acquisition Programs.

UNCLASSIFIED

To help understand the importance and limitation of human performance models in acquisition, let us consider the U.S Navy's goal of dramatically reducing on-board ship manning. One of the goals for the new Land Attack Destroyer (DD21) program is to reduce the manning from 300 to fewer than 100 sailors. This will require the concurrent design of new organizations, new equipment, and new tasks for Navy personnel. Further, design of the DD21 is a radically different design problem than engineering enhancements of legacy systems, which is the usual design process. Authentic models of organization and team behavior will be required early in the design process to optimize command structures by simulating the effectiveness of candidate structures in the planned mission space such as the new Land Attack mission of the DD-21 ship class. Further, as the design process proceeds, team models will be needed to support function allocation between crewmembers and automated components during the design of new systems.

There are fundamental differences between the designs of DD21 and legacy-based systems. This means that there will be a lack of analogies and historical data to support design decisions so team and organizational models will be critical for quantifying and evaluating the organizational, team, and task-teamwork design impacts of such radically different future systems. Unfortunately, team and organizational modeling is still in its infancy and even agreement on the psychological and organizational dimensions that need to be represented is lacking. Many researchers propose that teams perform additional tasks that groups of individuals do not, enabling them to (1) develop and maintain a shared situation awareness, (2) engage in compensatory and intra-team support behaviors, and (3) coordinate effectively taking into account the roles, relationships, and work load of teams members. Unfortunately, there is no clear evidence of what types of intermediate level representations and processes are required to emulate these behaviors and current team models are simply inadequate for the level of fidelity required by most military simulations.

Another key technology missing is a way to quickly allocate roles, translating design-oriented models into simulations that interact with external simulations of command and control centers. This capability would allow integrated product teams to readily evaluate designs from an operator's perspective, allowing them to predict performance levels as well as to produce preliminary data for planning training and manpower requirements.

UNCLASSIFIED

- As system designs are refined, human in the loop evaluations become increasingly prominent, involving ever more complex scenarios in rich, virtual environments. While these studies do assess the operators' ability to physically perform the desired tasks, the principal benefit is the assessment of the system to support the operator's perception and decision-making. In order for such an assessment to be meaningful, both simulations of BLUE and RED forces must present realistic behaviors that are not predictable yet are reproducible. In order to accomplish this level of fidelity, models of the following six operator characteristics from Table 1 will be required:

- mental ability
- learning and adaptation
- perceptual ability
- situation awareness
- problem solving
- decision making

Training

In most TTCP countries, training usually encompasses three sub-areas: instruction, training, and exercises. Instruction normally refers to classroom instruction that introduces new material to students. The material is usually presented by a lecturer or a computer and can cover a wide range of topics (e.g., anything from technical training on how to use a specific system to the procedures for Transfer of Authority). Training is an activity conducted to improve skills as well as to reduce knowledge-based decision making to rules or skills that may be executed quickly. Training often includes “hands-on” activities (e.g., using a part-task trainer) and is usually directed at teaching decision-making skills, reporting procedures, operational tactics, or other proficiencies. Exercises are formalized gatherings of a particular command echelon in which trained individuals practice their skills in a wider military context than encountered during training. Exercises can be conducted at strategic, operational, or tactical levels, often to evaluate operational plans or to evaluate the echelon's ability to execute its mission.

US Army leaders have often voiced the dictum that soldiers should train as they fight. Unfortunately, this is very difficult to achieve in practice. Even when weapons effects are simulated in live exercises, the human and physical resources consumed by these exercises make them prohibitively expensive for satisfying all Army training needs. A less expensive alternative

UNCLASSIFIED

is to train against a computer-generated force (CGF) using simulation. Current CGF models produce relatively predictable behavior tied closely to doctrine but have not been designed to perform as humans would. Current CGFs lack human emotions and personality traits that would be expected to produce believable variation in performance. The 1998 National Research Council report on Modeling Human and Organizational Behavior: Application to Military Simulations, as well as, several Blue Ribbon and Senior Navy management panels have identified that a shortcoming in the area of military simulations is the lack of behavioral realism in computer generated forces. Given the military's growing reliance on large-scale simulations as a means to prepare our warfighting teams, this deficiency has far reaching consequences. In particular, it is necessary to devise robust strategies for simulating higher-order processes in CGFs including decision-making, intent, deception, adaptability, creativity and problem solving.

Most US Navy training is team oriented, requiring trainees to develop important skills in coordination with other Navy team members. However, it is often inappropriate or inconvenient to train the entire team at the same time in the same instructional system. It is desirable to train a subset of the complete team (e.g., those who need additional training or who are available) without incurring the costs and problems of manning all team watchstations with suitably experienced personnel. The availability of behavioral and cognitive models of watchstation performance for all watchstations not being actively trained could substantially reduce cost while increasing flexibility and accessibility of simulation-based training.

The Department of National Defence in Canada has noted that weapons system operators will require high fidelity representation of human behavior to support simulation for training:

“Increasing pressure is being put on temporal and physical resources available for the conduct of individual and crew served weapons training. Weapons systems are being designed with ever increasing lethality and cost. Training for the proficient use of these systems implies a greater expenditure of effort to ensure crew and observer safety, training area environmental integrity and realistic target presentation. With the reduction in the number of training areas following the flurry of base

UNCLASSIFIED

closures with the end of the Cold War, it is becoming harder to find ample clean space for weapons training in existing bases.

Project L2624 Unit Weapons Trainer... will resolve the acknowledged capability deficiency by procuring simulation equipment for regular Land force Units and schools, militia unit armouries, militia training and support centers and Maritime command schools in support of individual and crew served weapons training for Canada's Land and Maritime Forces."

The requirements for this project and those similar can be met by providing training devices which incorporate realistic individual and crew human-simulator interfaces. This involves good representation of weapon feel in set-up, target acquisition and engagement. It will seldom be the case that target entities will be encountered in isolation or, as in the current range of operations, that targets are in close proximity to non-combatant actors. This implies a need for faithful representation of terrain and environmental factors, as well as target aspect and battlefield behaviors, both as an individually represented entity and as part of combat team. In training for situations other than war, the behavior of non-targets may have to be represented.

Human performance simulation has definite application to the design of high fidelity weapon system simulators and trainers. Psychophysical research must eventually be embedded in trainee/crew simulator interfaces. Portraying realistic target behavior in simulators leaves scope for applying intelligent agent technologies, to control single target action to the more sophisticated cooperative agent architectures for multiple target behaviors.

To explore the role and requirements for human performance models to support training, we first discuss the changing goals of exercises and then discuss the infrastructure that is used to conduct exercises, using the US Joint Simulation System (JSIMS) as an example.

For most of the TTCP countries, the primary purpose of exercises during the cold war era was to prepare their battle staffs to fight either Soviet or Soviet style armed forces. A typical exercise might, for example, bring together one or more command echelons to practice executing

UNCLASSIFIED

a variation of a war plan, with the primary training objectives being to practice the procedural skills associated with conducting a war.

With the end of the cold war, the range of military missions undertaken by TTCP countries has increased dramatically. TTCP countries are undertaking new classes of missions, such as humanitarian operations, peace enforcement, peacemaking, and peace building. During these new missions, the military undertakes non-traditional activities, interacting with a wide variety of cultures and people. Furthermore, such missions typically occur on short notice with no pre-existing deployment or contingency plans that consider the unique tensions of the geopolitical regions involved. Many of these missions require military command and control elements (C2) to work with non-traditional allies or non-government organizations.

This increased range of missions and actors has created a significant need for rapid, detailed concept development and experimentation. Many countries are looking to M&S technology to fill this need in order to study tactics, doctrine, command and force structures. For example, the US is developing JSIMS to provide a simulation environment with the fidelity necessary to evaluate and train for such new and complex missions.

JSIMS is being designed to be the next generation campaign level model for all US Services. The goal is to provide a simulation capability to support joint or U.S Service training, mission rehearsal, and education objectives. It is being designed to simulate the full range of military operations, covering all combat-arms services as well as support functions such as logistics, transportation, intelligence, medical, engineering, communications, and electronic warfare. Further, JSIMS will be able to simulate the impact of social, economic, and political factors, that affect or are affected by, military operations.

JSIMS includes a core simulation infrastructure consisting of simulation services (e.g., timing coordination services, network communication services, etc.) as well as mission space objects that include the physical environment and simulation elements. The physical environment includes representations of terrain, weather and oceanographic details. Simulation elements include both military and non-military entities such as weapons systems, C2 echelons, civilian population or non-government organizations.

UNCLASSIFIED

JSIMS is being developed through eleven cooperative, mutually interdependent acquisition programs. Individual Services, specialized agencies and commands have development programs that are creating specialized sub-sets of mission space elements. The complete set of elements is kept in a modeling and simulation repository, from which sub-sets are drawn to create tailored simulations for specific purposes.

The complete JSIMS is being created incrementally through a series of version releases. Each version provides some of the required functionality; successive versions add to the previously baselined functionality. The US Air Force is responsible for developing Air and Space entities through the National Air and Space Model (NASM) program. We will use this program as an example to discuss the range of human performance modeling needed to support training objectives.

The NASM Operational Requirement Document (ORD)⁴ and the Air and Space JSIMS Breakout Analysis Spread Sheet (JBASS)⁵ are two major documents that identify human performance model requirements for Air and Space domain entities within the JSIMS environment. The NASM ORD discusses both shortcomings of existing models and new requirements for NASM. While the shortcomings section does not discuss human performance models per se, it does discuss several general limitations of the model currently used by the Air Force, the Air Warfare Simulation (AWSIM). Many of these limitations can be tied directly to human modeling limitations.

For example, the NASM ORD notes that AWSIM has no provisions for modeling social, economic, or political factors across the range of military operations. Further, the ORD notes that AWSIM does not adequately model radar track identification problems, fratricide, visual range combat, or C2, all of which require high fidelity human performance models. Finally, the ORD states that AWSIM has no capability to represent information warfare; which also requires high fidelity models of individuals and C2 echelons.

⁴ Operational Requirements Document, HQ USAF 009-93-II, National Aerospace Warfare Model, 5 Jan 99.
http://www.afams.af.mil/doctlib/doctlib.cfm?AFAMS_FID_185

⁵ Air and Space (A&S) JSIMS Breakout Analysis Spreadsheets (JBASS), V 1.4.3., 3 Jul 2001.
http://www.afams.af.mil/doctlib/doctlib.cfm?AFAMS_FID_2614

UNCLASSIFIED

In the NASM requirements section, the ORD discusses specific human factors model requirements. These include adjustable levels of human functionality at the unit and national levels to account for effects due to training, military doctrine, morale, fatigue, national resolve, political influences, social and religious factors, and chemical or biological attack. The ORD states that NASM must automatically scale these human factors to represent the cumulative effects of a bombing campaign, they must have observable effects in the simulation through behaviors such as sortie generation time or aircraft and missile launch response times.

General NASM requirements, such as those listed above, are further refined in the Air and Space JBASS. The Air and Space JBASS identifies specific entities that the Air Force is responsible for creating in each JSIMS version. In this discussion, we will only consider the set of entities required at full operating capability. We will limit ourselves further by only considering the sub-set of these that are the C2 entities (i.e., we will not discuss aircraft entities).

There are 28 separate classes of entities that must be created, ranging from specific individuals, such as a Airborne Weapons Director who attempts to disambiguate radar tracks, to a fully-staffed Air Operation Center (AOC), employing hundreds of personnel. Typically, several variants of each entity must be created. Specific variations usually are created to represent operations in different theaters of operation (e.g., an AOC participating in a NATO mission versus an AOC employed in South West Asia), or different types of operations (e.g., war versus peace keeping). Further, each entity must have semi- and fully-automated versions (a semi-automated version is partially under human control), and must have high, medium and low fidelity representations. In all, the Air and Space JBASS calls for the development of over two hundred Air and Space C2 echelons.

The definitions of high, medium and low fidelity representations provide a great deal of insight into the challenges facing the JSIMS development teams. A low fidelity model is defined as one whose behavior is controlled via simple conditional logic. A medium fidelity model is one whose behavior is controlled by rule-based processing. Both the low and medium fidelity models behave linearly. A high fidelity model is one whose behavior is controlled by a non-linear, command-decision process. Very few, if any, contemporary models achieve the highest level of fidelity, yet this level of sophistication is often required to enable the users to achieve important

UNCLASSIFIED

goals such as the analysis of information operations or the psychological effect of a strategic bombing campaign on a civilian population.

The Air and Space JBASS is only one of the five domain BASSs. Further, we have only discussed a sub-set of the entities that are required to fully populate the domain. Extrapolating to the whole JSIMS enterprise, a minimum of a few thousand, high fidelity, human performance models will be required. Considering the performance requirements for JSIMS, one realizes that the low fidelity finite state machine models employed in the existing suite of simulations are inadequate to address the questions the military needs answered. No longer is the military primarily concerned with force on force attrition against Soviet style forces. The military's new requirements for simulation systems, as typified by the ORDs of the eleven JSIMS developmental teams, demand sophisticated capabilities that will enable them to model the whole range of human factors effecting individuals and organizations, complex decision-making (both in individuals and organizations), and the political, cultural, social, and economic factors effecting civilian populations.

Mission Rehearsal

Mission rehearsal differs from training, exercise, and instruction in that its goal is to evaluate and rehearse a *specific* mission in an environment that is as similar to the actual environment as possible. Further, mission rehearsal is usually conducted by a small group of individual units, such as a flight of aircraft or platoon of armored vehicles. Since mission rehearsal is specific to a set of conditions, it differs from training, which tends to be more generic.

Simulation-based mission rehearsal is a relatively new type of application arising from recent advances in simulation technology. Beginning in the late 1980s, a major effort was undertaken to create a digital battlefield that could help bring together a large number of separate simulators in a digital environment. Research was undertaken to develop networking and data interchange formats to link diverse simulators, and to create digitized physical environments in which the networked simulators would operate. In addition, research was initiated to create

UNCLASSIFIED

computer-generated forces, both semi- and fully-automated, that could participate in simulated missions and battles. These digitized battlefield environments are often referred to as networked virtual simulations since humans as well as virtual operators may operate simulated equipment in a simulated environment. Within such a networked virtual simulation environment, participants react as they would in a real mission using a workstation or simulator that use sophisticated multi-media technology and computer software to produce compelling views of the environment, including both friendly and opposing forces that behave realistically with appropriate weapons effects.

A networked virtual simulation environment could be used, for example, by a platoon to conduct an assault viewing the simulated terrain of the real world. This could enable them to assess fields of fire, evaluate cover and concealment possibilities, consider key terrain features, or observe various avenues of approach. In addition, the platoon would be able to realistically practice movement to the battle position as a platoon formation and the actions to be taken on contact with the enemy. Time permitting, the platoon might be able to rehearse and evaluate several different movement formations and courses of action.

A full, accurate, mission evaluation using network virtual simulation for rehearsal often requires judgment or constraints that could be provided by high fidelity human performance models instead of human participants. Such models would simulate the behavior of enemy as well as friendly collateral forces, especially when there are not enough human players available to represent all the required forces, or when the available players lack the training to accurately portray opposing forces. Human behavior models can also be used to simulate higher command and control echelon activities.

Because of the emphasis on evaluating the outcome of a specific mission, such models must have high fidelity behavioral representations across a wide range of variables. For example, it is important that the enemy decision making process be realistic, reflecting situational and environmental variables such as human factors (field of view, visibility, fatigue, etc.), current task demands, knowledge and training (i.e., how experienced is the opponent, in what tactics where they trained). It is also important that enemy unit movements are consistent, coordinated and reflect appropriate tactics and doctrine. Further, virtual participants may need to reflect

UNCLASSIFIED

behavior that an opponent might have realistically learned in previous (simulated) battles (Pew and Mavor, 1998).

The US Army has been the international leader in developing human behavioral representations for networked virtual simulations. These representations are usually not human performance models, *per se*, but semi-automated forces that typically couple a vehicle and behavioral representation together. The Army is usually concerned with the outcome of a large number of units interacting, and consequently, has focused on developing semi-automated forces that are not too computationally demanding; this focus in developing human and unit behavioral representations has meant the sacrifice of greater fidelity for speed.

The typical model framework used in networked virtual simulations is a finite state machine (FSM). FSM's include a list of states, a list of commands that can be accepted while in each state, a list of actions for each command, and a list of conditions required for an action to be triggered (Pew and Mavor, 1998). Usually the states of a finite state machine are derived from authoritative sources of doctrine, such as military task lists. FSMs do not model cognitive functions such as perception, information gathering, or information correlation. Because such models possess limited intelligence, they need to be controlled by a human who plans and executes their tactical actions, typically as an integrated unit rather than as system of coordinated components. Semi-automated forces work reasonably well for small unit tactics, but because they lack an underlying model of behavior it is difficult to scale them up to handle operation of larger size or scope (Ceranowicz, 1994; Downes-Martin, 1995; Pew and Mavor, 1998).

Several types of semi-automated forces have been developed to support networked virtual simulations. The US Army recently began the development of a more sophisticated type of semi-automated force (SAF) as part of the OneSAF program to replace some of the specialized SAF versions they currently use. This development will greatly expand the range of operations to which such models can be applied.

OneSAF is envisioned as a next-generation, computer-generated force that will represent the full range of operations, systems, and control process spanning those needed for the individual combatant and platform to the Battalion level. It is designed to work with WARSIM (Warfighter's Simulation), the US Army's component of the JSIMS enterprise, for higher-level

UNCLASSIFIED

organization and command. In addition, OneSAF is being designed to be interoperable with existing virtual network simulators as well as operational command and control equipment.

The OneSAF ORD⁶ states that it must be capable of simulating the missions and tasks of BLUE force light infantry, mechanized infantry and armor, artillery, air and missile defense, aviation, special operations, engineer, military intelligence, signals, military police, and NBC (nuclear, biological, and chemical) units. It must also be capable of simulating opposing force infantry, mechanized infantry and armor, combat support and combat service support units in the performance of their missions and tasks. All of these units must be represented at a sufficient level of fidelity to support analysis, mission rehearsal, and training objectives for a range of scenarios covering war, humanitarian aid and operations other than war.

The ORD describes several new classes of missions that the Army wants to be able to investigate and evaluate with OneSAF. These include Nation Assistance, Peace Making, Peace Building, Humanitarian Assistance, Arms Control, Counter Terrorism and Counter Drug Operations. In order to be able to accurately evaluate and train for such missions, the ORD calls for the new types of simulation capability. These include simulating behaviors associated with the local population, including those associated with displaced civilians refugees, mass migration and rioting. The ORD also calls for the simulation of non-government actors such as volunteer and international organizations. OneSAF simulation entities are to reflect differences in behaviors due to the nature and current state of the economy, the political structure of the society, the affect of military operations on society, and the reaction of the various societal groups to economic, military, political and information inputs from US forces and other societal groups.

Given these challenging requirements, it is unlikely that the traditional Finite State Model approach to modeling behavior will provide the required behavioral fidelity. It is expected that new classes of models will be required by the US Army to enable it to conduct the types of training, analysis and mission rehearsal envisioned to be prepared for future operations.

⁶ One Semi-Automated Forces (OneSAF) Operational Requirements Document (ORD), Version 1.1, 14 January 2000; Final draft.

Operations Research

Most TTCP countries have specialized groups of operations analysts that study and analyze future military requirements. In the US, the military services are developing the Joint Warfare System (JWARS) to model joint combat operations at the campaign and mission levels. JWARS is a constructive simulation that will provide multi-side representation of joint theater war. Example applications of JWARS include evaluating force, logistics, C4ISR⁷ adequacy, identifying resource shortfalls, developing force flows and sustainability requirements, and developing and evaluating war plans and supporting documents.

Historically, constructive models, such as JWARS, did not include explicit human behavioral representation such as models of human decision-making. Instead, these models normally used best-fit matching of library templates to the current situation reflecting subject matter expert opinion of what should happen.

It is uncertain whether the constructive models used by studies and analysis groups will ever incorporate human performance models. Discussions with CF military staff at the Army Simulation Centre in Kingston revealed that until HBR models can be validated, operational units would be unwilling to include them in analyses as they add even more uncertainty to a problem that is already difficult to evaluate. At the same time, HBR models would make useful additions to test the sensitivity of the results to factors suspected of affecting operator or system performance. More sophisticated models are needed because the relative strength of opposing forces, as represented in the Lanchester approach to modelling combat outcomes, is not what will “win the day”. Rather, information technology, in the forms of better command and control systems or the conduct of information operation missions, will play an increasingly crucial role in the conduct of successful operations. This change will require explicit and realistic modeling of human information processing because the human is the end user of information. Indeed, in information warfare, human information processing capability is often the target of the operation.

Stiner et al. (1994) note that the “*...post-Cold War strategic environment is ill-defined, dynamic, and unstable. The nature of this environment and the military threats it fosters indicate*

that US forces (1) will face a widely diverse range of adversaries equipped with an ever increasing array of sophisticated weapons, and (2) will require a span of operation and response capabilities that ranges from military operations other than war (OOTW) – such as deterring or engaging small, unsophisticated, fanatical terrorist groups – to conducting significant military operations against regional powers which may well possess advanced weapons systems, including nuclear, biological and/or chemical weapons of mass destruction.” Planning for such ill-defined scenarios is incredibly difficult and is hampered by the lack of empirical data. Extensive operations research may present the only viable method of exploring a sufficiently broad spectrum of events and contingency plans to have a significant chance of dealing with Stiner’s forecasted future that has since arrived. Consequently, the next generation of constructive models will require enhanced information processing abilities and incorporate additional aspects of human behavior to assess new military missions. This is even more necessary as analysts often have no past relevant experience to predict outcomes with any reasonable assurance. This need, however, is not currently stated explicitly in the requirement documents for these analysis tools.

Conclusion

There is a broad consensus among military simulation organizations that existing military simulations do not possess sufficient fidelity, particularly in the area of human performance, to answer the questions the military needs answered to be fully prepared for current and future operations. The new types of missions that the military is being asked to perform and the new types of operations (e.g., information warfare) that they are being asked to conduct cannot be adequately modeled with today’s technology.

At the campaign level, training simulations (e.g., JSIMS) have clearly defined requirements for high fidelity human performance models. These models are needed to define the operational concepts and training requirements for new classes of missions (peace keeping, information operations, etc.) and for dealing with new classes of actors (e.g., civilian

⁷ C4ISR – Command, Control, Communication, Computers, Intelligence, Surveillance, Reconnaissance

UNCLASSIFIED

populations, non-government organizations, et cetera). Constructive models for studies and analysis (e.g., JWARS), while not explicitly calling for human performance models, appear to have need for some human behavioral overlays, given the kind of studies anticipated. At the mission level, new models are clearly needed to support training objectives and mission rehearsal goals. In the past, simulations at this level have used fairly simplistic FSM representations rather than actual models of human behavior. Whether or not such FSM models can be scaled-up to answer the questions being posed by the today's military in the context of the Revolution in Military Affairs remains to be seen.

References

Andrews,D.H., Brown,J., Byrnes,J., Chang,J., Hartman,R., Irwin,J., Lightner,M., McQuay,W., Murashige,A., Sisti,A. & Wintermute,D. (1998) *Enabling technology: Analysis of categories with potential to support the use of modeling and simulation in the United States Air Force*. Final report of the Air Force Research Laboratory ad hoc working group on modeling and simulation enabling technologies, REPORT NUMBER: AFRL-HE-AZ-TR-1998-0041, PROJECT NUMBER: 1123 ACCESSION NO.: 99-00009, AFRL/HEOP, Brooks AFB.

Technology Investment Strategy For the Next Two Decades (2000) Retreived 7 January 2001 from http://www.crad.dnd.ca/public/tis/tis_e.html.

Anonymous (2000) TTCP HUM TP2 Strategic Review. In *Report of the Annual Meeting, Human Performance and Resources Group (HUM) Technical Panel HUM-TP2, Training Technology*. The Technical Cooperation Program. Niagara Falls, CA., 8-12May, Annex D, pp33-39.

Anonymous (1998) *NATO Modeling and Simulation Master Plan*. Document AC/323 (SGMS)D/2, Version 1.0. North Atlantic Treaty Organization. North Atlantic Council, Brussels. 7Aug1998.

Pew,R.W. & Mavor, A.S. (Eds) (1998) *Modeling Human and Organizational Behavior. Applications to military simulations.* Washington, DC: National Academy Press.

UNCLASSIFIED

Stiner, C., Fulwood, I., Stevens, S., Lutz, J., Steele, O., Crerar, S., Dorsey, M. M., Holmes, J., Jukoski, M., Nightingale, K., Coenen, B. & Davenport, S. (1994). *Report of the senior working group on military operations other than war (OOTW)*. (ARPA Order No. A119, Contract No. MDA972-93-C-0016): Advanced Research Projects Agency, Advanced Systems Technology Office.

Chapter 2. Assessment of Current Operational Capabilities in Human Performance Modeling

This chapter focuses on human performance modeling capabilities currently used in simulations for military training and analysis. Capabilities under development but not yet in the field or in use but not fully mature will be discussed in Chapter 3. This chapter is divided into two major sections. The first focuses on Computer Generated Forces (CGFs), reflecting current capabilities in the US and to a lesser extent, the UK and CA. The second section focuses on agents, and reflects current capabilities in AU.

At present, the military regularly employs a limited number of models to represent human action and decision making, representing both opposing and collateral, friendly forces in either virtual or constructive simulations. Prominent operational US systems that use these models include Modular Semi-Automated Forces, TacAir-SOAR, Joint Semi-Automated Forces, and Close Combat Tactical Trainer SAF.

Current Capabilities for Computer Generated Forces

ModSAF⁸

The most far-reaching computer generated forces system is the Modular Semi-Automated Forces (ModSAF) program. ModSAF and its variants are in use at military training sites as well as government and industry research facilities worldwide. The goal of the ModSAF program was to develop an architecture that would allow a small number of simulation operators to control higher echelon units and to enable the addition of various battlefield entities. The resulting ModSAF design is a fully distributable system, providing a repository of modules for developing simulations.

⁸ <http://www.modsaf.org/publicmodsaf1.html>

UNCLASSIFIED

The simulation of entities in ModSAF is divided into physical and behavioral models. Physical models are developed by defining and combining component models such as hull, turrets, weapons, and sensors. Behavior on the other hand is implemented via tasks. CGF software is implemented as library modules with strictly defined and documented public interfaces that can be combined to produce unique SAF systems. Small programs link these libraries together to form CGF applications.

The concept of a task is at the foundation of the ModSAF architecture (Calder, Smith, Courtemanche, Mar and Ceranowicz, 1993); tasks are behaviors performed by units or individuals. Within ModSAF, a task is represented by an entry in the persistent object (PO) database. Five types of tasks have been implemented in the ModSAF system: unit, vehicle, reactive, enabling and arbitration. Unit tasks define the actions of a unit; vehicle tasks define the basic behavior for single vehicles. These tasks have been represented in software via an asynchronous finite state machine. A FSM defines a set of states that represent the different lower-level actions making up the tasks, a set of inputs and outputs, and a set of transfer functions that cause transitions to new states when inputs to the state machine satisfy those functions. Task frames are used to group a collection of related tasks that run at the same time. A mission is a collection of task frames linked together to form a sequence. The ModSAF architecture supports the development of simple actions at the company level and below but any deviations from the list of states, commands, actions, and conditions included in the finite state machine result in unrealistic behavior.

Soar⁹

Soar is a computational programming architecture based on a proposed unified theory of cognition. Newell (1990) describes a cognitive architecture as the system of mechanisms that accesses encoded knowledge about the external world held in long-term memory in order to select actions to accomplish current goals. Soar's architectural structure consists of five characteristics:

UNCLASSIFIED

1. Soar formulates tasks in terms of problem spaces. All behavior in Soar is represented as movement through a problem space.
2. Long-term memory is constructed as a single production system. In Newell's terms, a production system is a form of pattern matching. The Soar production system was developed using Ops5 (Forgy, 1981)
3. Objects are defined by their set of attributes and values. At the simplest level, Soar's representation of objects is similar to object-oriented or schema representations. However, Soar does not use default values, attached procedures, or automatic inheritance structures.
4. Decisions about actions are communicated via preferences.
5. Soar uses a goal hierarchy with goals, subgoals, and alternative subgoals, to achieve performance. The Soar architecture divides knowledge into problem spaces and allows goals and actions in one problem space to be implemented via reasoning in another.

Since the early 1990's, Soar researchers have used the Soar architecture to develop intelligent forces for air missions including both fixed wing and rotary wing aircraft. A primary goal of the ModSAF system was its extensibility. The DARPA (Defense Advanced Research Projects Agency) WISSARD (What If Simulation System for Advanced Research and Development) project successfully demonstrated that software modules within ModSAF could be replaced. In the first demonstration of this feature, the Soar system successfully replaced the beyond visual range air-to-air combat tactical behavior provided through ModSAF's tasks frames with its own production rule system; the resulting application was called IFOR. TacAir-Soar, was constructed to support the training of navy F14 pilots in beyond visual range air-to-air combat (Jones, Tambe, Laird, and Rosenbloom, 1993; Rosenbloom, Johnson, Jones, Koss, Laird, Lehman, Rubinoff, Schwamb, and Tambe, 1994; Tambe, Johnson, Jones, Koss, Laird,

⁹ [<http://www.soartech.com>

UNCLASSIFIED

Rosenbloom and Schwamb, 1995)¹⁰. Soar Technology has extended its application to other air missions including: defensive combat air patrols, sweeps, close air support, interdiction, strategic attack, anti-armor, refueling, resupply (Laird, Johnson, Jones, Koss, Lehman, Nielsen, Rosenbloom, Rubinoff, Schwamb, Tambe, Van Dyke, van Lent, and Wray, 1995). The Soar architecture has been applied to the challenge of developing intelligent pilot agents for AH-64 Apache attack helicopters conducting an attrit mission¹¹. Soar agents are isolated from the underlying simulation environment; connectivity is provided through an interface to ModSAF. This interface enables access to sensor and weapon systems (Schwamb, Koss, and Keirsey, 1994). Soar Technology has recently applied its experience to the domain of dismounted infantry. As a proof of concept, they developed an Army Ranger special operations force team that performs a long-range reconnaissance patrol mission (Taylor, Koss, & Nielsen, 2001).

JSAF

Joint Semi-Automated Forces (JSAF) was developed by DARPA as part of its Synthetic Theater of War (STOW) project to support joint task force training on a large-scale battlefield. JSAF is currently one component of the JSAF Federation used by the J9 Directorate of the United States Joint Forces Command and the Navy Warfare Development Center for experimentation. As detailed in Ceranowicz, Nielsen and Koss (2000), JSAF is a descendent of SIMNET (Simulation Networking) SAF, ODIN SAF, ModSAF, and Soar. The JSAF architecture continues to use task frames for most behaviors – task frames are used to control units from individual combatants to battalions. The Soar agent architecture is currently used to control fixed wing aircraft through an interface to JSAF. Three different command agents were built under STOW funding: Army agents using constraint based satisfaction (Calder, Carreiro, Panagos, Vrablik, Wise, Chamberlain, Glasson, 1996); Army helicopter agents using Soar (Gratch, 1996); and Marine platoon agents using fuzzy rule sets (Goldman, 1996). Although these interfaces remain, they are not active in the current experiments. The present day version

¹⁰ Other references may be found at <http://www.soartech.com/htmlonly/publications.html>

UNCLASSIFIED

of JSAF continues to evolve through enhancements and extensions, however its dependence on task frames and production rules for behavior representation has not changed.

Close Combat Tactical Trainer

The Close Combat Tactical Trainer (CCTT) is the first of a family of networked simulators, including command, control and communications workstations, designed to portray the vehicles and weapons systems of a mechanized infantry or armor company team. CCTT¹² is a follow-on to SIMNET-T with more battlefield effects and a higher resolution terrain data base. Behavior is implemented in CCTT SAF through a finite state machine written in Ada. One of the key differences between CCTT SAF and ModSAF is CCTT's use of Combat Instruction Sets (CISs). CISs are derived from US Army Training and Evaluation Program Mission Training Plans and are a natural language description of tactical behavior. These CIS descriptions provided the sequence of actions that were captured in the CCTT SAF behavior code. Kraus, Franceschini, Tolley, Napravnik, Mullally, and Franceschini (1996) introduce key features of CISs, CCTT SAF, and ModSAF.

Australian Experiences with Agent Oriented Simulation

Agent Description

Australian activities in military simulation for Human-in-the-Loop (HIL) exercises and operations research are comparable in ambition to those of its TTCP partners. A broad, significant difference is the use of intelligent agents as representations of military operators. A decade of focus on simulating engagements and missions for

¹¹ <http://www.isi.edu/soar/soar-homepage.html>

¹² <http://www.stricom.army.mil/STRICOM/PM-CATT/CCTT/>

operations analysis with intelligent agents has developed a technology and methods in a way that the focus on providing computer-generated entities in HIL simulators did not.

The most significant work to date in terms of deployed applications lies in the air domain. The BattleModel¹³ software environment was developed to manage simulation agents, serving a similar function to HLA (High Level Architecture) or DIS (Distributed Interaction Simulation) supporting the requirements of a constructive-simulation based OR community.

Simulations embodying interactions between physical models and human reasoning capabilities, which are easy to explain and modify with Defence customers, have been implemented. New procedures using intelligent agents are linked to existing validated software (a huge investment) to define complex situations that involve tactical decision-making processes. The BattleModel represents a major increase in capability in facilitating the modelling and analysis of whole air missions with multiple aircraft types, roles, weapons, sensors, or communication systems.

An agent approach allows the analyst to work at a high level, formulating concepts and aims, while keeping the detailed computer programming hidden. The benefits of agent technology were demonstrated with the initial operational system SWARM leading to the proliferation of agent technology throughout the DSTO's (Defense Science and Technology Organization) Air Operations Division (AOD). The rapid uptake of agent technology by several different tasks forced the consideration of scalability challenges across the breadth of the system development process. Scalability issues were not isolated to the simulation infrastructure but occur at all levels in the software development process.

BattleModel is a simulation system that is capable of simulating the physics of whole air missions and the reasoning involved in such missions. This system provides DSTO with the ability to rapidly evaluate and test counter-air tactics for the RAAF (Royal Australian Air Force). It provides: high-fidelity simulation of combat aircraft,

¹³ <http://www.dsto.defence.gov.au/corporate/publicity/brochures/opianalysis.html>

UNCLASSIFIED

ground controlled intercept (GCI) controllers, and surveillance aircraft; advanced reasoning capabilities for modelling the pilot's decision making process; and, sophisticated visualisation tools to enable a better understanding of whole air missions. BattleModel allows DSTO analysts to rapidly create or modify tactics used as part of the pilot's reasoning process as well as quickly setting up simulation scenarios for operational studies.

Agent development within AOD has focussed almost exclusively upon a particular subclass of intelligent agents – BDI or Belief-Desire-Intention agents. The Belief-Desire-Intention model of rational agency has resulted in several language implementations, from the research languages such as JAM, IRMA, and C-PRS, to the industrial systems such as dMARS (Distributed Multi-agent Reasoning System) and JACK (JAVA Agent Compiler and Kernel).

Efforts are taken to ensure the language independence of application design. Not only is this good software engineering practice but it also mitigates risk in reliance on a single source. An appreciation of the importance of implementation language independent descriptions, designs and documentation was gained during the transition from the intelligent agent language dMARS to its likely successor JACK.

dMARS (the Distributed Multi-Agent Reasoning System) was a language developed by the Australian AI Institute (AAII) as an industrial implementation in C++ of PRS (The Procedural Reasoning System) developed at SRI. JACK¹⁴ (JAVA Agent Compiler and Kernel) is seen as the likely successor to dMARS. Developed by Agent Oriented Software (AOS) Pty Ltd it offers a similar set of functionality within a lightweight JAVA implementation. To date JACK has been used by AOD for a single simulation development as a technology test-bed but it is being actively researched and developed in cooperation with AOS.

A decision about which is the most useful implementation follows from the requirements analysis for each of the projects and, as can be clearly seen in Table 2.

¹⁴ Jack is downloadable from www.agent-software.com

UNCLASSIFIED

Another implementation of the BDI agent architecture, Attitude, used within DSTO's Information Technology Division. To date Attitude is not industrially supported and has the status of a research language.

Table 2. Summary of some DSTO Intelligent Agent Applications.

Application	Number of Agents	Purpose	Length of typical scenario	Geographic size of typical scenario	Size of biggest team	HIL Version
AEW&C	24	Acquisition Tactics Development	2 Hours	500nm	6	YES
VAE	8	Training	20 Minutes	500nm	2	YES
F-18	8	System Upgrade and Tactics Development	5 Minutes	100nm	4	PROTOTYPE
F-111	4	Tactics Development	20 Minutes	200nm	2	NO
P3	6	Tactics Development	4 Hours	800nm	6	YES
Helicopter Operations	4	Acquisition	30 Minutes	60nm	4	NO

Representation of Team Structures

The representation of the behavior of aggregates and issues of command and control, are being studied in multi-agent systems and in military simulation. There are applications of team and organisation models to many domains; the manufacturing industry in particular is increasingly aware of the capacity for intelligent agents to improve processes that require coordination or some kind of collaborative activity. The development of STEAM for the synthetic Theatre of War by Tambe et al. (1995) is a good example of military applications. A state of the art summary is to be found in the

proceedings of the 1999 IJCAI¹⁵ (International Joint Conference on Artificial Intelligence) workshop on Team Behaviour and Plan Recognition¹⁶.

The addition of team and organisation modelling constructs to existing agent systems is currently underway as a part of an ongoing process of improving the functionality of agent languages. In the military context, this allows for the development of a military mission planning tool with an ability to predict mission outcomes in response to combat activities. The modelling must accommodate dynamic changes forced on the Command and Control structure, such as battle casualties.

AOD's simulation development, which incorporates Intelligent Agents (IAs) to account for human decision-making processes, allows the military planner to assess the capability of different force options and to develop operational tactics for a range of air platforms and their associated systems and weapons. However, team structures are "hard-wired" in code with no capability for rapid or dynamic changes other than by extensive additional coding. Moving IA technology out of the laboratory environment should provide enhanced and more timely Command Decision aids to military planners. The transitioning process requires the development of a model of military teamwork, with software architectures supporting dynamic role allocation, team tactics, and team structures.

The requirements in this work are that

- C2 structures and individual agent behaviours will be specified separately
- C2 structures, team tactics and roles need not be replicated in individual agents
- *teamed agents* will have in-built ability to allocate or re-allocate roles

The major benefits are that

- simulations can be developed, scaled-up and modified easily

¹⁵ <http://ijcai.org/>

¹⁶ Papers may be downloaded from <http://www.agent-software.com.au/shared/demosNdocs/teams1.html>

UNCLASSIFIED

- explicit representation of C2 structures and team tactics will support automated recognition of team intentions in observed enemy behaviour

Teamed agents require

- existing agent technology to be extended to include team concepts
- each agent to have an internal view of the organisational structure, roles and assignments

Teamed agents will enable detailed evaluation of

- team structures
- role assignments
- force deployment options
- roles and responsibilities within teams

Teamed agents will provide a mission planning and decision support environment. Initial applications will be operational effectiveness studies in areas such as air combat with both fixed, including F/A-18, Airborne Early Warning and Control, and rotary wing platforms. Maritime while more complex Land applications could follow.

Agent Oriented Design for Simulators

Agent Oriented design is based on mental states: intentions, plans, desires, goals, roles, beliefs and descriptions of the situations in which agents act. It complements existing object oriented, procedural and structured methods, integrating with or extending existing methods and technology.

Agent Oriented concepts embody the principles of abstraction and encapsulation that are a standard way for software engineers to manage the complexity of complex

UNCLASSIFIED

systems development. These approaches help defence customers and military operators understand the process and details of creating representations of human operators.

According to the requirements of the activity, a number of architectures are employed. The requirements analysis and design process is undertaken in an “agent-oriented” fashion that is independent of the subsequent implementation. Locally developed techniques allow for the integration of intelligent agents with legacy and third party software and simulators. This includes the capacity to link via standard protocols such as HLA to the wider distributed simulation community. Table 2 indicates some of the characteristics of a range of DSTO developed intelligent agents simulations.

REFERENCES

Calder,R., Carreiro,R., Panagos,J., Vrablik,R., Wise,B., Chamberlain,F. & Glasson,D. (1996). Architecture of a command forces command entity. *Proceedings of the 6th Conference on Computer Generated Forces and Behavioral Representation*, pp. 19-30, Orlando, FL.

Calder,R., Smith, J., Courtemanche, A., Mar, J. & Ceranowicz, A. (1993) ModSAF Behavioral Simulation and Control. *Proceedings of the 3rd Conference on Computer Generated Forces and Behavioral Representation*, pp. 347-358, Orlando, FL.

Ceranowicz,A., Nielsen,P. & Koss,F., (2000) Behavioral representation in JSAF. *Proceedings of the 9th Conference on Computer Generated Forces and Behavioral Representation*, pp. 501-512, Orlando, FL.

Forgy,C L (1981) *OPS5 User's manual*. Technical report (July). Computer Science Department, Carnegie-Mellon, Pittsburgh, PA

Gratch,J., (1996) Task decomposition planning for command decision making. *Proceedings of the 6th Conference on Computer Generated Forces and Behavioral Representation*, pp. 37-48, Orlando, FL.

UNCLASSIFIED

Goldman, S., (1996) Knowledge acquisition and delivery: constructing intelligent software agents. *Proceedings of the 6th Conference on Computer Generated Forces and Behavioral Representation*, pp. 31-36, Orlando, FL.

Jones, R., Tambe, M., Laird, J. & Rosenbloom, P. (1993) Intelligent automated agents for flight training simulators. *Proceedings of the 3rd Conference on Computer Generated Forces and Behavioral Representation*, pp. 33-44, Orlando, FL.

Kraus, M., Franceschini, D., Tolley, T., Napravnik, L., Mullally, D. & Franceschini, R. (1996) CCTT SAF and ModSAF behavior integration techniques. *Proceedings of the 6th Conference on Computer Generated Forces and Behavioral Representation*, pp. 159-170, Orlando, FL.

Laird, J., Johnson, L., Jones, R., Koss, F., Lehman, J., Nielsen, P., Rosenbloom, P., Rubinoff, R., Schwamb, K., Tambe, M., Van Dyke, J., van Lent, M. & Wray, R. (1995) Simulated intelligent forces for air: The Soar/IFOR project 1995. *Proceedings of the 5th Conference on Computer Generated Forces and Behavioral Representation*, pp. 27-36, Orlando, FL.

Newell, A. (1990) *Unified theories of cognition*. Cambridge, MA: Harvard University Press.

Rosenbloom, P., Johnson, L., Jones, R., Koss, F., Laird, J., Lehman, J., Rubinoff, R., Schwamb, K. & Tambe, M. (1994) Intelligent automated agents for tactical air simulation: A progress report. *Proceedings of the 4th Conference on Computer Generated Forces and Behavioral Representation*, pp. 69-78, Orlando, FL.

Schwamb, K., Koss, F. & Keirsey, D. (1994) Working with ModSAF: Interfaces for Programs and users. *Proceedings of the 4th Conference on Computer Generated Forces and Behavioral Representation*, pp. 395-400, Orlando, FL.

Tambe, M., Johnson, L., Jones, R., Koss, F., Laird, J., Rosenbloom, P. & Schwamb, K. (1995) Intelligent agents for interactive simulation environments. *AI Magazine*, 16(1), 15-40.

UNCLASSIFIED

Tambe,M., Schwamb,K., Rosenbloom,P., (1995) Building intelligent pilots for simulated rotary wing aircraft. *Proceedings of the 5th Conference on Computer Generated Forces and Behavioral Representation*, pp. 39-44, Orlando, FL.

Taylor,G., Koss,F., Nielsen,P., (2001). Special operations forces IFORs. *Proceedings of the 10th Conference on Computer Generated Forces and Behavioral Representation*, pp.301-306. Norfolk, VA.

Chapter 3 Human Performance Research and Development Projects

This chapter summarizes on-going work across the TTCP nations aimed at enhancing the competence, realism and military application potential of human performance models. This work extends from basic science to applied research and development. Representative publications are cited in association with each program considered.

United States

United States Army

Human Behavioral Modeling Program

The US Army Research Institute (ARI), with the help of commercial and academic leaders in modeling and simulation, are exploring ways to make CGF behave more like humans, making them susceptible to combat fatigue, giving them rudimentary personality traits, and allowing them to react realistically to combat events. ARI's first effort in this area was a joint project with Science Applications International Corporation (SAIC). This project developed mathematical models that account for sleep deprivation effects, circadian rhythm, experience, and aggressiveness. These models have been tested in conjunction with two different simulations and shown to affect command decisions and simulated combat outcomes.

ARI is seeking to make CGF act more realistically by representing basic emotions, personality factors, stressors, and training into human performance models. As a result, the soldiers training with simulations will encounter a wider range of plausible situations allowing them to practice their combat skills against a realistic foe. Including these factors makes CGF less predictable and more like a real opponent because the reactions of CGF to combat events vary as a function of emotions and personality as well.

UNCLASSIFIED

as cognitive factors. Soldiers trained against an intelligent unpredictable foe learn adaptive behaviors and resourcefulness reducing the risk of cognitive rigidity.

ARI is currently supporting efforts to further humanize CGF by integrating human emotions and personal traits into human performance models. These efforts are organized by industrial partners and a brief synopsis describing each approach to this goal is presented below.

Related Publications

Gillis,P.D., Hursh,S., Guest,M.A., Sweetman,B. & Ehrlich,J.A. (In press). *Cognitive behaviors for computer generated command entities*. (ARI Technical Report). Alexandria, VA: U. S. Army Research Institute for the Behavioral and Social Sciences.

Gillis,P.D. & Hursh,S. (1999) Using behavior moderators to influence CGF command entity effectiveness and performance. *Proceedings of the 8th Conference on Computer Generated Forces and Behavioral Representation*, pp. 237-251, Orlando, FL.

***SOAR Technology*¹⁷**

Researchers at Soar Technology, Inc. are integrating their rule-based architecture with a connectionist model of emotions. The connectionist model assumes that emotions arise from a combination of pleasure or pain, arousal, attention and time components. The selected application incorporates emotions and individual differences into the behavior models of synthetic virtual helicopter pilots in a battlefield simulation. The pleasure or pain system interprets the level to which a stimulus represents a threat or enhancement to survival. In turn, pleasure and pain stimulate the arousal system. Different personality types may be more or less susceptible to events that generate arousal, pleasure or pain.

¹⁷ <http://www.soartech.com/>

UNCLASSIFIED

These personality differences lead to distinctive decision making profiles that can produce crucial performance differences in combat situations.

The model postulates that clarity and confusion are also important in determining behavior. A confused person is less likely to respond so as to minimize pain or maximize pleasure. Emotional attributes combine with deliberate cognitive processes and background knowledge in working memory to generate strategies, reasoning and external behavior. Simultaneously, the cognitive model evaluates the environment and status of internal goals (situational awareness). The connectionist model uses this information in computing new values for each emotional attribute. Hence, the emotional attribute values vary during combat to influence combat behavior dynamically.

Related Publications

Chown,E. (1993) *Consolidation and Learning: A Connectionist Model of Human Credit Assignment*. Doctoral Dissertation, The University of Michigan.

Rosenbloom,P., Laird,J. & Newell,A. (1993) *The Soar papers: Research on integrated intelligence*. Cambridge, MA: MIT Press.

***Natural Selection Inc.*¹⁸**

Natural Selection Inc. (NSI) is using Evolutionary Programming techniques to develop a realistic, non-rule-based, intelligently-interactive combat simulation for training two or more combat teams of various skill and intelligence levels. Evolutionary Programming evolves combat plans and behaviors based on their ability to satisfy some stated goal or condition. For example, the goal might be to minimize own force casualties or to maximize damage to selected military targets. To improve the fidelity of their approach, NSI is also modeling personal traits of individual decision-makers, such as loyalty, risk-taking propensity, motivation, and social ability. For demonstration purposes, NSI is interfacing the evolutionary programming modules with a version of the

constructive-simulation JANUS. NSI selected four basic mission types to demonstrate the effectiveness of evolutionary programming techniques in generating intelligent and realistic behaviors: (1) attack a predefined position, (2) attack a mobile enemy, (3) defend a predefined position, and (4) defend from enemy attack (withdraw).

In phase 1 of this effort, NSI used a basic attack scenario as a proof-of-concept demonstration. In that scenario, six Red Force tanks were defending a strategically important landmark. The goal of the Evolutionary Programming controlled Blue Team was to attack the position defended by the Red Team and secure the objective, surviving at all costs. This technique successfully evolved plans that out-manoeuvred the Red Team to reach the objective.

While the primary focus of phase 2 is to incorporate personality and other soft factors into the model, the goals of phase 2 include developing a human performance model that can be used immediately by the military to improve the realism of their simulations. The outcome of this effort will be an intelligently-interactive human performance model that is adaptable to virtually any simulation engine, includes personal traits and allows interaction or comparison with humans. Recent work has focused on including a higher meta-learning algorithm (with memory) in the model.

Related Publications

Porto,W.V., Fogel,D.B. & Fogel,L.J. (1998) Generating novel tactics through evolutionary computation. *SIGART Bulletin* (Fall), pp.8-14.

Fogel, D.B. (2000) What is evolutionary computation? *IEEE Spectrum*, pp.26-32, February.

¹⁸ <http://www.natural-selection.com/>

***Psychometrix Associates, Inc.*¹⁹**

Psychometrix Associates, Inc. selects individual differences based on empirical evidence indicating a high potential to affect soldier and commander performance, modeling them as parameters of a cognitive architecture developed by Psychometrix. A large number of individual differences and affective factors (e.g. ability factors, risk tolerance, anxiety or stress tolerance, aggressiveness, fear, mood) are being considered for inclusion in the model. These factors change the parameter values in the cognitive architecture that, in turn, change the computer-generated agents' behavior. The cognitive architecture will also incorporate individual goals and expectations, working memory, and a model of attention. In the Psychometrix approach, goals are selected based on the current emotional state; they in turn trigger approaches that guide perceptual and cognitive activities and decision-making. A simulation test bed environment is being designed to demonstrate and evaluate the architecture's performance. The test bed will display the goals and expectations, emotional states, and decisions made by the intelligent agent being modeled. The test bed will map all of the possible situations and behaviors, as well as the agents' choices within its environment, providing a complete record of the agents' behaviors and states within the context of the combat scenarios.

Related Publications

Hudlicka,E., Zacharias,G. & Psotka,J. (2000) Increasing realism in human agents by modeling individual differences: Methodology, architecture, and test bed. *Proceedings of the AAAI Fall Symposium*. Falmouth, MA.

Hudlicka,E. (2000) Integrating Model of Trait and State Influences on User Behavior. *Proceedings of the 2nd Workshop on Attitude, Personality, and Emotions in User-Adapted Interaction*. Sonthofen, Germany.

¹⁹ <http://www.psychometrixassociates.com/>

Micro Analysis and Design²⁰

Micro Analysis and Design is attempting to improve the realism of computer-generated force (CGF) entities in constructive simulations, developing a set of algorithms and data structures for including variables such as aptitude, training, and stressor effects that can be integrated with other types of available software packages used for developing human performance models.

The work is currently progressing on three main thrusts. One thrust is focused on developing the aptitude algorithms, learning curves, and stressor algorithms that will eventually influence the performance variables in the human performance models. Work in support of this thrust included a review of the literature on individual and team learning theories, military requirements for human performance modeling and a search of the literature for empirical data that describes the effects of training on human performance. Work toward developing the learning curves has also included the development of a data collection questionnaire for obtaining estimates from soldiers on how their training affected their proficiency in combat performance. This questionnaire was administered to platoon leaders and platoon sergeants from Armor divisions at Fort Riley, Kansas and Fort Carson Colorado. Data from the questionnaires has been analyzed and used to develop learning curves for classroom, simulator, and field training effects.

The second thrust of the project is the development of a software tool that will allow a user to enter information about the training and aptitude of a population of operators in a human performance model. This information will be used as input to the learning curves to calculate the appropriate changes to performance variables in the human performance models. In order to make the tool generalizable for any type of training or human performance model, it will be completely configurable by the user. This feature of the software tool will allow users that have data to generate their own learning curves and use those learning curve algorithms to affect the performance variables in the models. A part of this effort was the development of a test bed model to

²⁰ <http://www.maad.com/>

UNCLASSIFIED

test and demonstrate the functionality of the software tool in correctly modifying performance variables.

The third major thrust of the project is to develop the architecture for communicating performance variable values between the software tool and the human performance models. Included in this effort is the selection of an appropriate entity-based constructive simulation such as ModSAF, OneSAF, or JSAT to apply the tool described above. At this time, the eventual target platform is OneSAF. However, since all three of these simulations share the same code base, this work began on the JSAT software. One of the major challenges in developing the communication architecture is in identifying the performance and potential stressor variables in the simulation that can be modified.

Related Publications

Archer,R., Walters,B., Oster,A., & Goldberg, S. L. (2000). Improving Soldier Factors in Human Performance Models, *Proceedings of the 2001 Interservice/Industry Training, Simulation and Education Conference*. pp. 271-279, Orlando FL.

Oster,A., Walters,B., Archer,R., & Edgin T.(2001). Mathematical Algorithms for Training Effects Determination in CGF. *Proceedings of the 2001 Interservice/Industry Training, Simulation and Education Conference*. pp. 281-289,. Orlando FL.

Institute for Creative Technologies²¹

The primary goal of the Institute for Creative Technologies (ICT) is to explore techniques to create highly realistic training simulations for the Army that rely on virtual reality, artificial intelligence, sound, and emotions. ICT has three projects exploring immersive multimedia and AI (Artificial Intelligence): pre-visualization project for future combat systems, experience learning system, and mission rehearsal.

²¹ <http://www.ict.usc.edu>

UNCLASSIFIED

The mission rehearsal exercise (MRE) project seeks to create a training environment that will challenge soldiers to practice their decision-making skills under more realistic circumstances. The simulation is populated with virtual humans within the context of a civil unrest scenario. The behavior of these virtual humans was controlled either by scripts, an AI reasoner, or an AI reasoner augmented by an emotional model. The emotional behavior component of the MRE relies on the Physical Focus model of the Interactive Pedagogical Drama system (Marsella, Johnson, & LaBore, 2000; Gratch & Marsella, 2001). Boston Dynamics created bodies for the virtual humans; Haptek added expression to the faces. This project is still in its early stages. At present formal evaluations of the MRE as a learning environment are planned for next year.

Related Publications

Gratch, J. & Marsella, S. (2001). Modeling emotions in the mission rehearsal exercise.

Proceedings of the 10th Conference on Computer Generated Forces and Behavioral Representation, pp. 457-465, Norfolk, VA.

Marsella; S., Johnson; W.L., & LaBore; C. (2000). Interactive pedagogical drama.

Proceedings of the 4th International Conference on Autonomous Agents, pp. 301-308.

The United States Navy and Marine Corps

In 1999 ONR (Office of Naval Research) and NAWCTSD (Naval Air Warfare Center Training Systems Division) launched a comprehensive research and development initiative on Human Behavioral Representation (HBR) and Computer Generated Forces (CGFs) for Military Simulation to investigate emerging technologies and methods for cognitive modeling and behavioral representation. This project's objective is to develop sound scientific and technological bases for CGFs to be used in military simulations at varying levels of aggregation from individual combatants to Naval task forces. Specific goals being pursued to meet this objective include creating CGFs that exhibit realistic

UNCLASSIFIED

behaviors in scenario-specific reasoning, learning, problem solving, chain-of-command strategy development and implementation, motivation and planning.

Results and guidelines from this research will bring us closer to realizing fully automated, realistically behaving simulated forces that will provide worthy adversaries and intelligent friends in virtual and constructive simulations for training, mission planning and rehearsal, analysis, acquisition and command decision aiding.

Basic (6.1) Research

This initiative includes basic, applied and advanced R&D programs. The basic (6.1) research of the integrated program is performed under the Realistic Modeling of Human Behavior component of a larger ONR 6.1 Program in Modeling and Simulation managed collaboratively by ONR and NAWCTSD. The objectives of this 6.1 program focus on HBR and CGF architecture development and studies. Specific projects currently funded are described below.

***Robustness in Behavioral Modeling*²²**

Soar Technology, Inc. has undertaken the task of improving the robustness of the TacAir-Soar behavioral model for computer-generated forces with the idea that the problem of brittleness is fundamentally a knowledge problem that arises out of ignorance, not representation. The claim is that brittleness is caused by a lack of knowledge about what should be expected. This project began in June 2000 and progress to date has focused on planning methods to best assure that a complex behavioral generator consisting of thousands of rules can be hardened for field conditions while still retaining its flexibility for research and experimentation.

Soar Technology plans to approach a solution to this problem through a number of distinct tasks aimed at common failures. First, one common failure in TacAir-Soar is a “state no-change.” This arises when there is insufficient information in the goal state to

²² Principal Investigator: Paul Nielsen, Soar Technology

UNCLASSIFIED

reach a conclusion and no further information is available to create a sub state that can resolve the missing information. An analysis of updating to Soar 8 style operator termination has shown that many instances of this type of failure are preventable. Upgrading to this newer Soar version would ensure that any operator is retracted when its applicability conditions no longer exist.

By far the most common failure in TacAir-Soar can be characterized as human error. Information that should be provided to the system is incorrect or incomplete. Eventually this will be caught in testing, but often not until that point in the mission when the information is actually needed. To handle such cases Soar will include further error checking in the Exercise Editor, add redundant error checking to mission planning, add explicit expectations of preconditions and add additional rules that act as assertions on the allowable values of attributes.

To more actively find system failures they will undertake to incorporate qualitative reasoning about the physical characteristics of behaviors. That will provide a predictive model of errors that may occur in the future so that the agents may begin to take action to prevent them, and allow mental simulation to decide whether a given course of action is feasible before it is actually attempted.

Extending COGNET CGF Human Modeling Capabilities²³

The purpose of this research is to investigate methods and techniques to improve technology for modeling human performance for a range of military applications, emphasizing the integrated representation of cognitive, perceptual, and motor performance. The application of principal interest is computer-generated forces (CGFs) used in distributed simulations, and with emphasis on CGF application to naval forces. The military significance derives from the resulting availability of a toolset and framework for human behavioral representation that will be highly usable and efficient. This toolset will support Navy M&S applications for stand-alone training, embedded training and mission rehearsal, as well as system evaluation, intelligent interfaces and

²³ Principal Investigators: Wayne Zachary and Floyd Glenn, CHI Systems

UNCLASSIFIED

decision support. The project is adding performance-modeling capabilities to the integrated cognitive architecture system called COGNET (Cognition as a Network of Tasks), in the areas identified and discussed in the 1998 National Research Council study on Modeling Human and Organizational Behavior. The specific capabilities most extensively addressed are

- sensory and motor abilities and constraints
- mechanisms to represent individual differences and situational effects such as stress, fatigue, effects of extreme environments, et cetera
- agent self-awareness (physical and cognitive states) that affects metacognitive processes, increasing adaptation, improving error or interruption recovery, and producing cooperative behaviors
- flexibility and granularity-independence in representation
- computational efficiency of the simulation engine
- the usability of the system
- interoperability of the models with broader simulation frameworks such as HLA

The primary approach to accomplishing these goals has been to embody the research results in a software tool called CGF-COGNET. This is an extension of the original COGNET cognitive modeling approach. CGF-COGNET permits flexible granularity in component process representation using efficient, top-down development of computer generated force models to meet behavioral representation requirements listed above.

CGF-COGNET is being applied in the follow three projects

1. SCOTT (Synthetic Cognition for Operational Team Training), a Navy program to develop synthetic teammates for a deployable training system.

UNCLASSIFIED

2. AMBR (Agent-based Modeling and Behavior Research), an Air Force program to develop and demonstrate realistic human performance models of air traffic control, using the HLA standard
3. MTRS (Mission-Training and Rehearsal System), an Air Force ACC (Aerospace Combat Command) program to provide human performance modeling and intelligent tutoring in a mission training and rehearsal system for Air Force Ground Theater Air Control System (GTACS) C2.

Representative publications:

Eilbert,J., Le Mentec,J.-C., Glenn,F., Zachary,W. & Weiland,M. (2000) Cognitive models with accurate representation of human sensory and motor action behavior for simulating command environments. *Proceedings of SPIE 45th Annual Meeting*.

Zachary, W. W. & Le Mentec, J.-C. (2000). Modeling and simulating cooperation and teamwork. In M. Chinni (Ed.), *Military, Government, and Aerospace Simulation*, 32(3), 145-150. San Diego, CA: Society for Computer Simulation International.

Zachary, W. W. & Le Mentec, J.-C. (2000). Incorporating metacognitive capabilities in synthetic cognition systems. *Proceedings of the 9th Conference on Computer Generated Forces and Behavioral Representation*, pp. 513-521, Orlando, FL.

Zachary, W. & Le Mentec, J.-C. (1999). A framework for developing intelligent agents based on human information processing architecture. *Proceedings of the IASTED International Conference on Artificial Intelligence and Soft Computing*. pp.427-431. Anaheim: IASTED/Acta Press.

Le Mentec, J.-C., Glenn, F., Zachary, F. & Eilbert, J. (1999). Representing human sensory and motor action behavior in a cognitive modeling architecture. *Proceedings of the 8th Conference on Computer Generated Forces and Behavioral Representation*, pp. 587-593, Orlando, FL.

UNCLASSIFIED

A Hybrid Architecture for Human Performance Modeling²⁴

The objective is to answer two questions: How can high-level cognition be efficiently grounded in interaction with the environment? How can a cognitive-based computer generated agent be made more adaptive to the environment? The strategy is to combine reactive learning and a cognitive model. The hypothesis is that an integration of these two approaches will create a system that combines the best of both reactivity and high-level cognition (e.g. planning), with learning both at the reactivity level and at the cognitive level.

The overall goal is to build a cognitive model that uses machine-learning methods to improve the performance of the model over time, making it more adaptive. An empirical approach will be used to demonstrate the utility of combining these systems. The study will proceed along two fronts using a Micro-Air Vehicle (MAV) control task that allows reactivity and planning level cognition complexity to be varied. An ACT-R (Adaptive Control of Thought-Rational) cognitive model will be designed, based on human subject performance, dealing with high-level planning issues. Human participants will have to perform the task with increasing levels of both planning difficulty and higher levels of reactivity. A learning agent (Samuel) will be constructed to perform the task reactively and comparisons made between the ACT-R and learning agent performances.

Representative publications:

Trafton, G., Schultz, A., Bugajska, M., Gittens, S. & Mintz, F. (2001). A hybrid cognitive and reactive architecture. *Proceedings of the 10th Conference on Computer Generated Forces and Behavioral Representation*. pp. 267-276, Norfolk, VA.

Schultz, A. & Trafton, G. (2000). Learning in a hybrid cognitive and reactive architecture. *AAAI Symposium on Parallel Cognition for Embodied Agents*. pp. 37-38.

²⁴ Principal Investigators: Alan Schultz and Greg Trafton, Naval Research Laboratory

Extending ACT-RPM to Multitasking Environment²⁵

The purpose of the project is to challenge and extend cognitive modeling of human interaction with complex environments with an ultimate objective of making cognitive modeling a practical tool for rapid development of predictive models or intelligent agents for complex, dynamic environments. The goals of this project are to apply the ACT-R²⁶ hybrid cognitive architectures to complex dynamic environments, determine which features perform well in such situations and develop new mechanisms for those features that do not scale well.

The first phase of the project modelled multi-tasking behavior using ACT-R in an HLA simulation of air traffic control. Although the model is quite simple, it successfully accounted for a broad range of subject data, including aggregate performance, individual differences in performance, response times, choice percentages, and both amount and type of errors. The ability to detect events and interrupt current tasks in favour of new, more urgent tasks was added to ACT-R's perceptual module, reproducing subjects' multi-tasking behaviors in this goal-centred architecture. A definition of cognitive workload based in existing architectural primitives was also added, and predicted workload values were very similar to subjective ratings.

The second phase of the project, now on going, is aimed at modeling the impact of learning and teamwork on human performance. Learning provides a difficult challenge for cognitive models, but also provides insight into the effects of instruction, training and the origins of individual differences. The intent is to model learning behavior into a variant of the task used for the first phase of the project to account for differences between novice and expert performance. Teamwork is an essential component of today's work environments but is often neglected in cognitive modeling approaches. Teamwork will be captured as an emergent property of interaction between independent cognitive agents. Preliminary work to date has included the release of an extension of ACT-R that allows communication between multiple ACT-R models in a highly efficient manner, and

²⁵ Principal Investigators: Christain Lebiere and John Anderson, CMU

²⁶ <http://act.psy.cmu.edu/>

UNCLASSIFIED

the development of a newer form of knowledge compilation to model the acquisition of instructions.

Potential applications of this research are numerous. A detailed computational account of multi-tasking and cognitive workload might lead to the development of algorithms based on predictive cognitive models that would allocate tasks dynamically to optimize the workload between several human operators. A better understanding of the process by which instructions and practice are assimilated into procedural fluency might lead to the design of better instructional material and practice regimen. A quantitative account of the development of teamwork could help guide the design of environments that foster its emergence.

Representative publications:

Lebriere, C, Anderson, J. R. & Bothell, D. (2000). Integrating working memory and perception in a dynamic environment using the ACT-R architecture. *Symposium on Human Performance Modeling Evaluation at the XIVth Triennial Congress of the International Ergonomics Association and the 44th Annual Meeting of the Human Factors and Ergonomics Society*. Santa Monica, CA: Human Factors Society.

Lebriere, C, Anderson, J.R., & Bothell, D. (2001). Multi-tasking and cognitive workload in an ACT-R model of a simplified air traffic control task. *Proceedings of the 10th Conference on Computer Generated Forces and Behavior Representation*. pp. 91-98, Norfolk, VA.

Graphical Visualization of Situational Awareness for Computer Generated Forces²⁷

A working prototype of the Situational Awareness Panel (SAP), integrated with agents developed within Soar and using JSAT Task Frames, has been developed demonstrating the capability to provide an architecture-independent tool for visualizing

²⁷ Principal Investigator: Randy Jones, Soar Technology, Inc

UNCLASSIFIED

the internal reasoning of simulation agents, a potentially powerful tool for training, trouble-shooting and model validation. The instantiated SAP focuses on modeling the pilot of a fixed wing combat fighter, and includes a visual representation of the pilot's current understanding of the combat situation, the goal hierarchy currently driving his behavior, and a running history of significant remembered events leading up to the current moment in time in the combat scenario.

A draft Applications Programmer Interface (API) has been developed that should allow future developers and users to tailor the SAP to work to their own applications. The API specifies an interface for passing information from intelligent agents into the SAP, as well as software patterns for presenting the information graphically. The API relies on a software component design, to maximize modularity and minimize dependence on particular behavior architectures, graphical presentations and task domains. Because the tool is designed for use with intelligent agents, it includes components that represent specific types of knowledge, including goals, operators and alternatives. This is very similar to a GOMS (Goals-Operators-Methods-Selection Rules) level of abstraction, providing enough detail to focus on representations of human behavior while not being too closely tied to any particular behavior architecture.

With the basic API design and prototype completed, the current focus is on more advanced features of the SAP. Work continues to improve the usability of the SAP, incorporating useful features based on feedback from end users. In addition, high-level functions for knowledge traceability, log-and-replay, and question answering are under development. Traceability provides a method to follow and record the process of making decisions by ACT-R agents. Log-and-replay allows the use of the SAP study details of events that occur during a simulation. Question answering allows users to explore an agent's selection of goals, means of satisfying goals and alternative courses of action. Each of these advanced features relies on the component-oriented knowledge representation defined by the SAP API. Components allow the modular use of the SAP's knowledge representation not only for graphical presentation, but to support these advanced features. Most recently, efforts have focused on refining the API to make sure it supports these advanced features effectively. The final objective is to provide useful

documentation, so that developers will be able to use and adapt the SAP to their applications.

Testing an interface to explain cognitive models to experts²⁸

The objective of the SAP project is to improve the Situation Awareness Panel (SAP), supporting the development and use of SOAR. The SAP explains the situation knowledge and behavior of Soar models of agents to trainees, analysts and onlookers, such as domain experts verifying the behavior of the Soar agents. In December 2000, the team at Penn State's School of Information Sciences and Technology started three main tasks to improve the SAP's usability. These are to:

- evaluate the SAP using expert evaluation and task analysis
- evaluate the SAP using local HCI (Human-Computer Interface) experts
- evaluate the SAP using non-local domain experts

Developing Computational Models of Naturalistic Decision Making²⁹

This is a three-year effort to develop a computational representation of Recognition-Primed Decision Making (RPD). The goal of this work is to improve the behavioral realism of computer generated forces by introducing theoretically sound representations of human decision-making. The nod to theory here is not just a means to an end, but rather an integral aspect of a computational process model that not only stands to improve realism in CGFs, but also has the reciprocal potential to deepen our understanding of human decision making. Toward this end, Micro Analysis and Design (MAAD) has worked closely with Klein Associates to understand the RPD model theory and to build on this understanding as implementation issues arise. MAAD has constructed a test bed model along with a variety of tools to help understand and evaluate the performance of the RPD computational model. Although the analysis of the test bed

²⁸ Principal Investigator: Frank Ritter, Penn State University

²⁹ Principal Investigators: Ron Laughery, Micro Analysis and Design & Gary Kein, Klein Associates

UNCLASSIFIED

is ongoing, evaluation of the model in a second, more complex environment is planned that will examine conflict detection in an air traffic control (ATC) environment. The current model architecture will also be modified to capture the effects of an agent dynamically re-evaluating the salience of cues during the decision making process. The architecture that comes of this research should result in models of decision-making that can be developed quickly and efficiently in a variety of contexts.

Representative publications:

Archer, S., Warwick, W. & Oster, A. (2000). Current efforts to model human decision making in a military environment. *Proceedings of the Military, Government, and Aerospace Simulation Symposium*. pp. 151-155.

Research into the Mathematical Modeling of Human Behavior³⁰

This research program applies soft computing (SC) techniques to modeling human behavior of high level cognitive tasks. SC is the development of hybrid systems using fuzzy logic, artificial neural networks, probabilistic reasoning, among others, to represent intelligent systems such as human behavior. Probabilistic reasoning includes genetic algorithms and other forms of evolutionary computation, as well as Bayesian approaches, imprecise probabilities and many other areas.

In this project, SC techniques will be applied to high-level cognitive activity in the military air C2 domain. This domain was selected for several reasons:

- involves high level cognitive processing where there are multiple goals and tasks
- COGNET models of similar problems exist that can be used to assess the added value of SC
- NAWCTSD has a C2 simulator that will be used for human C2 studies

³⁰ Principal Investigator: Frank Cardullo, SUNY, Bingham

UNCLASSIFIED

- The higher-level cognitive modeling has direct application to the DD-21 and CVX programs in reduction of ship personnel.
- SC has application to the reduction of role players in large-scale war games such as JSIMS

The unique architecture involves three main divisions integrated into a hybrid model of human behavior: a heuristic branch, a data driven branch and a branch based on first principles of physics and human perception. The heuristic branch employs fuzzy and classical logic as well as signal processing as tools, incorporating doctrine, complex procedures and risk assessment as representative human behaviors. The risk assessment mode involves modeling the human's assessment of the relative risk to the operator's domain of the various entities in the problem space. The data driven branch employs neuro-fuzzy models, genetic algorithms and perhaps the optimal control model as tools, incorporating test data from the air warfare controller simulation. The human perception branch employs classical modeling tools using theoretical principles, psychophysical limits as well as eye and head tracking.

The architecture and the risk assessment module have been designed and the coding is about to begin. The optimal control model is coded and has been tested with manual control data. The test bed has been implemented and debugged. A method for integrating the three branches of our architecture is being developed. The test protocols are being designed and will be submitted to the human subjects research review committee for approval. These experimental data will be used in developing the models in the data driven branch.

Research on Autonomous Synthetic Entities³¹

The thrust of this research is to expand our understanding of what is required to quickly develop human-like computer generated forces that have low computational

³¹ Principal Investigator: John Laird, University of Michigan

UNCLASSIFIED

overhead. Development of tools and methods for fast development and evaluation of CGFs in Soar is a central part of this project. Key features of the project are:

- Low computational overhead: This work has led to development a “lite” version of Soar and an associated interface that have significant less computational requirements compared to the standard Soar for computer generated forces. Analysis of the possibility of developing architectures for CGFs that can modify their modeling accuracy based on the available computational resources is also underway.
- Quick development: Tools are being developed that learn by observing a human perform the task. These tools will also allow one to quickly develop “personalized” synthetic entities and emphasize the specific characteristics of human behavior. A tool is under development that compares human behavior to an existing synthetic entity with the goal of supporting automatic validation and identification of errors. Further, it could identify human errors during training. Initial tools in these areas have been created and are currently being extended to address behaviors that are more complex and noisy data sets.
- Human-like behavior: A significant justification for the learning by observation work is to capture human-like behavior. Also under development is a new CGF capability, anticipation. Finally, a method is being developed for applying the Turing Test to synthetic entities with some testing to determine key characteristics for producing human-like performance.

Representative publications:

Laird, J.E. (2000). It knows what you’re going to do: Adding anticipation to a Quakebot. *Proceedings of the AAAI 2000 Spring Symposium Series: Artificial Intelligence and Interactive Entertainment*. pp. 55-61.

Laird, J.E. & van Lent, M. (2000). Interactive computer games: Human-level AI’s killer application. *Proceedings of the National Conference on Artificial Intelligence, AAAI*

Applied (6.2) Research

The applied R&D component of this integrated program is Computer Generated Forces for Scenario Based Training contained in the 6.2 Instructional Technology Program between ONR and NAWCTSD. The objectives of this 6.2 program are to develop and investigate the feasibility of instructional strategies using HBR and CGFs in Naval training systems.

Team Training with Synthetic Teammates (SYNTHERs)

It is important to identify what characteristics such watchstation cognitive simulations must have in order to provide effective simulation based training and to demonstrate that those characteristics are achievable with current cognitive modeling technology. The SYNTHetic teammembERs (SYNTHERs) project will explore current and developing cognitive modeling technologies, applying them in a team-training environment. The objective is to produce intelligent agents to act as replacements for missing human participants. Simulated teammates are a promising alternative to human teammates, because they are always available, may be modeled after experienced training personnel, and may be more cost effective in the long run. A SYNTHERs-like solution will be needed if the Navy is to achieve anytime-anywhere (24x7x365) team training in an affordable manner. Two current studies are described below.

Study of Behaviors for Synthetic Teammates in Team Training³²

This initiative explores how to overcome the bottlenecks of the availability and drawbacks of human teammates for training teams in synthetic environments, while keeping the advantages of human participants: the opportunity to learn in a collaborative and cooperative fashion. The research challenge lies in keeping the advantages of collaborative and cooperative behavior typically associated with human teammates in simulated teammates. This program will review the relevant available research data, and

³² Principal Investigators: Alma Schaafstal, TNO, and Denise Lyons, Naval Air Warfare Center Training Systems Division

UNCLASSIFIED

will explore how intelligent teammates should be defined and modeled include the possibility of cooperative learning, as well as optimizing individual and team learning experiences.

Synthetic Force Applications of Cognitive Modeling Technology in Navy Training Systems³³

The goal of this research is to investigate means of improving the fidelity of synthetic forces to be used in affordable Navy training systems, particularly team training. This research focuses on improving the Navy's ability to affordably train individuals and teams in their essential teamwork skills and other job requirements using synthetic team members. The research is using CHI Systems' COGNET modeling formalism and the iGEN™ cognitive agent software toolset. This project uses an enhanced iGEN™ toolset developed to improve the behavioral realism of CGFs and adding meta-cognitive capabilities including an increased sense of self-awareness.

Work to date has focused on enabling SYNTHERs to represent and apply communications network (“net”) discipline and to maintain enhanced situation awareness of other team members’ speech interactions when interacting with human warfighters. With sufficiently capable speech technologies for recognition and generation of speech, these advances will allow SYNTHERs to interact with human warfighters in a much more realistic way than previously possible. Three SYNTHERs are being constructed in this effort (Electronic Warfare Supervisor (EWS), Anti-Air Warfare Commander (AAWC), IDS). The first (EWS) is completed and the second and third are nearing completion. Current efforts are further exploring means of incorporating COTS (commercial-off-the-shelf) speech technologies (e.g., IBM’s ViaVoice and Microsoft’s Speech API) to take advantage of these new approaches in the SYNTHERs project as well as other Navy R&D efforts.

Representative publications:

³³ Principal Investigators: Wayne Zachery and Jim Hicinbothom, CHI Systems Inc.

UNCLASSIFIED

Hicinbothom, J. & Lyons, D. (1999). Synthetic forces for team training. *Proceedings of the 8th Conference on Computer Generated Forces and Behavioral Representation*. pp. 333-338, Orlando, FL.

Stokes, J. (in Press). Speech interaction and human behavior representations (HBRs). *Proceedings of the 10th Conference on Computer Generated Forces and Behavior Representation*. pp. 467-476, Norfolk, VA.

Hicinbothom, J. (in Press). Maintaining situation awareness in synthetic team members. *Proceedings of the 10th Conference on Computer Generated Forces and Behavior Representation*. pp. 231-241, Norfolk, VA.

Composable Behaviors in JSAF³⁴

Micro Analysis & Design is developing a Composable Behavior Server (CBS) to provide an easy and quick method to compose behaviors for JSAF entities. The Aegis Combat System on a United States cruiser, the CG-65 (USS Chosin), is being modeled as a demonstration of the CBS.

A Micro Saint task network model of Aegis that includes Anti-Air Warfare (AAW) and Surface Warfare (SuW) was developed then reviewed by subject matter experts (SMEs) at NAWCTSD. The model will reside within the CBS and will be used to provide Aegis target engagement and weapon selection parameters in real time to JSAF.

TADMUS scenarios, centered on an Aegis Cruiser, have been designed to stress the Aegis system operations and will be replicated to serve as test cases. An additional JSAF scenario involving an Aegis equipped ship has also been developed including interaction with other friendly, enemy and neutral entities. HLA will be used to pass data between the CBS and JSAF, requiring interaction additions to the JSAF FOM (Federation Object Model) from the RPR (Real-time Platform Reference)-FOM.

³⁴ Principal Investigator: Steven Peters, Micro Analysis & Design

UNCLASSIFIED

Cognitive and Analytic Modeling Approaches for Tutoring Systems (CAATS)³⁵

The purpose of this research is to investigate the capability of different types of human behavior representations, varying along a continuum of cognitive fidelity, to support effective trainee feedback in a dynamic training environment. Cognitive fidelity is defined as the extent to which the HBRs are believed to capture the underlying human cognitive processes associated with interacting with the environment.

Study 1 investigates the relationship between HBR cognitive fidelity and the effectiveness of the feedback generated to increase our understanding of HBR-feedback requirements in a dynamic training system, and thus improve our capability to develop effective automated training systems. Study 2 will focus on the following variables: (a) the level of cognitive fidelity of the HBR used to generate the feedback, (b) the level of granularity of the feedback, and (c) the format in which the feedback is presented to the trainees.

To date, the CAATS project has investigated three modeling techniques (fuzzy logic, classification and regression trees, and COGNET), and has constructed expert models of performance on the GT-ASP (Georgia Tech Aegis Simulation Program) task. In addition, the corresponding instructional materials for all three HBR approaches have been developed.

Representative publications:

Bolton, A.E., Holness, D.O., Buff, W.L. & Campbell, G.E. (2001). The application of mathematical models in training systems: A viable approach to cognitive modeling? *Proceedings of the 10th Conference on Computer Generated Forces and Behavioral Representations*. pp. 497-505, Norfolk, VA.

³⁵ Principal Investigators: Amy Bolton, Wendi Buff, & Gwendolyn E. Campbell, Naval Air Warfare Center Training Systems Division

Advanced (6.3) Research and Development

The advanced (6.3) research and development of the integrated CGF program is focused on demonstrating and measuring the effectiveness of HBR and CGFs in prototype Naval Training Simulations. There are two on-going projects, described below.

Intelligent Agents to Support Real-Time Exercise Control and Data Collection in Distributed Training Exercises³⁶

The objectives of this project are to develop, implement, and demonstrate instructionally sound training strategies including automatic run-time scenario adaptation and life-cycle management for large-scale modeling and simulation exercises. Tools such as software-based intelligent agents, common database structures, human performance modeling, and performance measurement systems will be incorporated within these strategies. Work will focus on streamlining exercise planning in support of real-time control to accommodate student needs and review of training performance.

Several research tasks are involved, supporting the development of distributed simulation training:

Task 1: Determine technical and functional system requirements.

Task 2: Develop data structures to support training management.

Task 3: Develop an instructional expert module to drive instructional decision making.

Task 4: Develop intelligent software processes to support automated and semi-automated management of the distributed exercise life cycle.

Task 5: Develop graphical user interfaces (GUI) for exercise planners, controllers, analysts, and observers.

³⁶ Principal Investigators: Randall Oser, Michael R. McCluskey, Elizabeth Blickensderfer, Gwendolyn E. Campbell, and Denise M. Lyons, Naval Air Warfare Center Training Systems Division

UNCLASSIFIED

Task 6: Conduct system testing, tuning, and validation of distributed exercise management using inputs from the joint training community and human-in-the-loop experiments.

The initial phases of the research have been completed for the identification of technical and functional requirements for the training system, identification of data sources and structures to support training management functions, preliminary architecture for the intelligent software processes, and initial development work for the graphic user interfaces. Continued development in each of these task areas is dependent upon the completion of cognitive task analyses with trainers, training managers, exercise controllers and observers. A work plan and a comprehensive set of detailed questions have been completed for the cognitive task analysis. Data from the cognitive task analyses will permit completion of the requirement and data structure definitions, followed by intensive software development for the intelligent agents. This development activity will also involve the completion of the instructional expert module to guide information exchange and instruction.

Representative publications:

Oser,R., McCluskey,M.R., Blickensderfer,E., Campbell,G.E. & Lyons,D.M. (2000) Real-Time Modification of Large Scale Exercises: Supporting the Management of Human Trainer Resources. *Proceedings of the 2000 Interservice/Industry Training, Simulation and Education Conference*, pp. 725-734, Orlando FL.

Synthetic Cognition for Operational Team Training (SCOTT³⁷)

SCOTT will apply advanced human behavioral representation methods in delivering a prototype intelligent, stand-alone training delivery system for deployed aviation personnel. This system will enable an individual or team to practice crucial advanced team skills on an anytime, anywhere basis with simulated teammates serving in roles of missing team members. The overall objectives of the project are to:

UNCLASSIFIED

- Develop an Aviation Team Training (ATT) research testbed
- Develop synthetic teammates that can function in realistic mission simulations
- Develop capabilities for automated scenario generation, performance measurement, and diagnostic feedback, which incorporate human performance modeling and advanced voice recognition system technology
- Develop a low-cost personal computer-based graphical user interface (GUI) networked to the testbed to provide the operator a menu-driven scenario generator and a replay capability for debriefing purposes, and
- Integrate, test and demonstrate the full system

The near-term project goal is a human-in-the-loop feasibility demonstration of an intelligent simulated teammate performing in the role of a Combat Information Center Officer (CICO) in the E-2 Hawkeye 2000 aircraft. The simulated teammate will be developed using the CGF-COGNET system and iGENTM software. It will perform the tasks of the CICO in a strike warfare mission scenario and in concert with a human trainee (an E-2 Air Control Officer - ACO), and a simulated E-2 Radar Officer (RO). The initial scenario and demonstration will feature a constrained battlespace to include E-2 operator interactions with simulated JSAT-TacAir Soar entities (i.e., F-14s). The capabilities of the synthetic teammate to interact with a human trainee and external simulated entities are keys to program success and will receive major emphasis in the early development phase. Subsystem frameworks comprising essential instructional support capabilities for the prototype training system, including automated performance assessment, on-line feedback, and after action review, also will be implemented in the initial development work.

Working groups convene regularly via teleconferences to coordinate work across several key development areas. A detailed scenario, including communication scripts, of a deep inland strike mission involving numerous aviation assets and controlling agencies

³⁷ Principal Investigator: Wayne Zachary, CHI Systems

UNCLASSIFIED

has been developed with E-2 SMEs at the Naval Strike and Air Warfare Command. Conversion of the E-TRACS software programs and interfaces to support the ATT testbed is continuing and voice system requirements have been specified. Hardware for the testbed has been acquired and an initial lab configuration to support current software versions of TacAir Soar and JSAT established. TacAir Soar and JSAT have been implemented in the testbed along with a terrain model to support the demonstration scenario. Requirements to introduce SoarSpeak into the testbed and develop the grammars to support the scenario have been defined.

Representative publications:

Zachary, W., Santarelli, T., Lyons, D., Bergondy, M., & Johnston, J. (2001). Using a community of intelligent synthetic entities to support operational team training. *Proceedings of the 10th Conference on Computer Generated Forces and Behavioral Representation*, pp. 215-223, Norfolk, VA.

United States Air Force

Combat Automation Requirements Testbed (CART) Program³⁸

The Air Force Research Laboratory's (AFRL) Human Effectiveness Directorate initiated the Combat Automation Requirements Testbed (CART) program to develop and demonstrate a human performance modeling environment that will provide a means for modelers and analysts to represent human decision making and tactics in a model that can interact with other simulations using the HLA. The top-level technical objectives are to: (1) develop tools that enable creation of realistic operator models based on goal orientation, (2) develop methods that connect operator models to a joint constructive environment, (3) demonstrate the technology via Case Studies using warfighter domains, and (4) transition the technology to the warfighter, acquisition and industry.

³⁸ Principal Investigator: Dave Hoagland, AFRL

UNCLASSIFIED

The CART human modeling environment provides users with an indication of the effect that operator performance has on system lethality and survivability. The hierarchy of goal-to-task is key to CART's translation of real-world actions and events into useable operator models. CART permits the user to decompose a mission into high-level goals (perform mission, threat evasion, attacking a target, etc). Once goals are established, the user can create a high fidelity human performance model (HPM) that includes detailed operator tasks that support each higher-level goal. Creation of lower level tasks in the model can be based on real world experience, engineering analysis, interviews with subject matter experts, or simply assumptions about how the operator will interact with the crew interface or the external environment.

Case studies were conducted to compare CART model behavior with that of real pilots (HIL) for similar scenarios and tasks. Both HPM and eight experienced pilots performed six repetitions within similar air-to-ground time critical target (TCT) scenarios. All of the 36 targets were destroyed by the HPM while pilots destroyed 47 targets out of 48 trials, or about 98%. A congruency ratio, measuring the overall correlation of the HPM and HIL performance, was calculated to be 0.7799 [Hoagland et al 2001a]. The square of this ratio provides an indication of the similarity in the variability of the two performances; in this case 0.61.

A univariate analysis was performed to determine if a statistical difference existed for eight dependent variables (DVs) across all scenarios *without* consideration of correlations. Three of the eight DVs had a significant difference between the HPM and the HIL ($p < 0.05$). However, the strength of the effects in the population (using the eta-square index) was low to moderate (less than 0.70). A multivariate repeated measures analysis was also completed in order to account for the *correlation between DVs*. This analysis provides a more accurate depiction of the differences between the HPM and HIL conditions. Given that there were significant ($p < 0.05$) correlations among the dependent variables taken as a whole (HPM and HIL combined), the true test of the relationship between independent variable (model vs pilot) and the eight DVs required a doubly multivariate analysis. There were three sets of DVs wherein each variable was correlated significantly with each other across all six scenarios, thus resulting in three doubly multivariate models. All three models demonstrated fairly strong differences (eta-squares

UNCLASSIFIED

of approximately 0.70 or greater) between the HPM and HIL. However, *post-hoc* analysis revealed that significant differences ($p < 0.05$) occurred only for two of the eight DVs and these differences were associated with outliers for certain pilots and scenarios.

The CART team developed the Case Study 1 model with the use of subject matter experts from the JSF (Joint Strike Fighter) program office and the SIMAF who had previously employed a limited set of tactics used in a prior VSWE exercise. CART model testing was initially completed using these tactics and results revealed a limitation in system effectiveness (low probability of finding and ultimately destroying ground targets). With these results in hand, the team quickly developed a *new tactic* using the same crew interface, but instead with an emphasis on the coordinated use of the cockpit's SAR (Synthetic Aperture Radar) and TIR (thermal infra-red) displays together. The HPM predicted that the new tactic would result in a much higher probability of detection, identification, and destruction of targets for all six scenarios. Later, during our pilot trials, subjects were trained on the new tactic and *actual performance was found to be higher than previously experienced* using the old tactic - and also was closely aligned with HPM results. The implication of this discovery is tremendous for the warfighter as CART can be employed as a tool not only to develop, analyze, and establish traceable crew system requirements, but also as a means to optimize tactics of current and conceptual aircraft systems.

The results thus far on the CART program lead us to the conclusion that our goal-oriented operator model matched pilot performance with good strength. The results of this project support the CART vision to provide the Air Force with the capability to maximize total human-system performance while saving time and money in acquisition using realistic human models.

References:

Hoagland,D.G., Martin,E.A. & Brett, B.E. (1999) The Combat Automation Requirements Testbed (CART) Program: Improving the DoD's Requirements Process Through Inclusion of Realistic Operator Performance. *Proceedings of the Simulation Interoperability Workshop*, 14-19 March 1999, Paper Number 99S-SIW-129.

UNCLASSIFIED

Hoagland,D.G., Martin,E.A., Brett,B.E., Doyal,J.A., LaVine, N.D. & Sargent R.A. (2000) The Combat Automation Requirements Testbed (CART) Program: Results and Lessons Learned from Recent Testing of Advanced Human Performance Models Interacting with DoD Constructive Simulation. *Proceedings of the Simulation Interoperability Workshop*, 17-22 September. Paper Number 00F-SIW-023.

Hoagland,D.G., Martin,E.A., Anesgart,M., Brett,B.E., LaVine,N.D. & Archer,S. (2001a) Representing Goal-Oriented Human Performance in Constructive Simulations: Validation of a Model Performing Complex Time-Critical-Target Missions. *Proceedings of the Simulation Interoperability Workshop*, 25-30 March, Orlando,FL. Paper Number 01S-SIW-137.

Hoagland,D.G., Martin,E.A., Anesgart,M., Brett,B.E., LaVine,N.D. & Archer,S. (2001b) Combat Automation Requirements Testbed (CART) Case Study One: Results from Modeling Human Behavior Within High Fidelity Constructive Simulations. *Proceedings of the 10th Conference on Computer Generated Forces and Behavioral Representation*, pp. 425-435, Norfolk, VA.

Agent-based Modeling and Behavior Representation (AMBR) Role Player Intelligent Controller Node (RPICN³⁹)

Command post exercises (CPXs) are formalized gatherings of a particular command echelon whose purpose is to train warfighters on the use of real-world plans, procedures, and equipment for C2 of Air Force forces. Exercises may be conducted at strategic, operational, and tactical levels and are used both to evaluate existing operational plans and to develop and test new operational concepts and tactics. At present, the Air Force uses the Air Warfare Simulation (AWSIM) as the simulation engine for most exercises.

³⁹ Principal Investigator: Lt. Michael Dooley, AFRL

UNCLASSIFIED

The AMBR-RPICN program is using intelligent agent technologies and behavioral models to partially automate role player functions to enhance role player interaction with AWSIM. The RPICN is a computer program consisting of an easy to use interface and a set of intelligent agents. The agents contain rule sets that embody standard response behavior to known model limitations. The agents monitor events occurring in the exercise, watching for problems. When a problem is detected, the agents alert the role player and provide them the ability to correct several model limitations with the click of a button. The RPICN thus frees the role player from having to parse through pages of scrolling text to identify problems and partially automates the required response behavior. It addition, it improves role player efficiency by allowing them to take on additional responsibilities and it reduces the overall number of role players required to run a training exercise. The RPICN program recently finished software development and will be tested at the Air Force Agency for Modeling and Simulation (AFAMS) in Orlando, FL on July 10th and 11th, 2001.

Agent-based Modeling and Behavior Representation (AMBR) Model Comparison⁴⁰

The Air Force Research Laboratory's Human Effectiveness Directorate started a research initiative called the Agent-based Modeling and Behavior Representation (AMBR) Model Comparison Project as part of an investment strategy for increasing the realism of human behavior models in defense simulations. One of the primary goals of the AMBR Project is to advance the state of the art in cognitive and behavioral modeling for military applications.

The AMBR Model Comparison attempts to accomplish this by organizing a series of comparisons among alternative modeling approaches. There will be four comparisons (or "rounds"), each overseen by a neutral moderator. The first two rounds are complete. The modeling goal in the first round was multi-tasking, and the task domain required a simplified version of en-route air traffic control. Modelers using ACT-R, D-COG, EPIC (Executive Process Interactive Control)-Soar, and CGF-COGNET participated in Round

⁴⁰ Principal Investigator: Kevin A. Gluck, AFRL/HEAS

UNCLASSIFIED

1. All were able to approximate the trends and central tendencies of the data, despite differing approaches to multi-tasking capability within these architectures.

In Round 2, DMSO sponsored the conversion of the simulation environment and models into a persistent, HLA-compliant human behavior representation testbed. All four models were modified to be compliant with HLA. Most human participants reported no perceptible difference between the HLA and non-HLA versions of the environment, and with few exceptions, the models produced similar results in both versions. DMSO's Federation Development and Execution Process (FEDEP) was found to be flexible enough to support federations involving human behavior representation.

In Rounds 3 and 4, the modeling focus will be category learning. The task for Round III will retain the multi-tasking perceptual-motor features of the simplified air traffic control task used in Round 1 and described in detail in Pew, Tenney, Deutsch, Spector and Benyo (2000) with one modification: an embedded category learning task will replace the speed query. Multiple aircraft will query the controller (the one that is being modeled) about the possibility of changing altitude. The controller will make a decision to authorize an altitude change based on a multi-dimensional attribute matrix that might include dimensions like aircraft size, level of atmospheric turbulence and current altitude. The Controller must learn the appropriate responses based on feedback received through the user interface resulting from their decisions. This concept learning task is based on the original laboratory study by Shepard, Hovland and Jenkins (1961), and modeling studies reported by Nosofsky et al. (1994). The Round 4 task will be designed to further stress the models and examine their capabilities based on the results of the Round 3 model evaluations. We anticipate a focus on the ability of models to adapt from one set of learned concepts to a new, changed set of concepts based on the same or a similar set of concept attributes. Other manipulations such as the workload of the perceptual motor task may also be explored as deemed appropriate given the results of Round 3.

Representative publications

UNCLASSIFIED

Gluck, K. A., & Pew, R. W. (2001). Lessons learned and future directions for the AMBR model comparison project. *Proceedings of the 10th Conference on Computer Generated Forces and Behavior Representation*, pp. 3-6, Norfolk, VA.

Nosofsky, R.M., Gluck, M.A., Palmeri, T.J., McKinley, S.C., & Gauthier, P. (1994). Comparing models of rule-based classification learning: a replication of Shepard, Hovland, and Jenkins (1961). *Memory & Cognition*, 22, 352-369.

Tenney, Y. J., & Spector, S. L. (2001). Comparisons of HBR Models with Human-in-the-loop Performance in a Simplified Air Traffic Control Simulation with and without HLA Protocols: Task Simulation, Human Data and Results. *Proceedings of the 10th Conference on Computer Generated Forces and Behavior Representation*, pp. 15-26, Norfolk, VA.

Intelligent Mission Controller Node (IMCN⁴¹)

The Air Force trains the AOC staff and the Joint Forces Air Combat Commander (JFACC) by having them accomplish the Air Tasking Order (ATO) planning and execution cycle as part of a command post exercise (CPX).

The intelligent mission controller node (IMCN) program is using intelligent agents and rule-based inferencing technology to link the Theater Battle Management Core System (TBMCS) to several Air Force simulations, including the current simulation used to support CPXs (AWSIM) and the new simulation environment under development (NASM). TBMCS is the main Air Force system used to task air assets against targets. The system is used to create and disseminate the Air Tasking Order (ATO) and the Airspace Control Order (ACO).

TBMCS creates a series of data fields containing mission details when it creates an ATO. The IMCN program is addressing the problem of mismatch between the data fields of TBMCS and the receiving simulations. In real world operations, several of the

⁴¹ Principal Investigator Dave Perme, Breakaway Solutions, Inc

UNCLASSIFIED

data fields are filled in by the Wing Operation Center (WOC) responsible for executing the mission, the Package or Mission Commander who will command the overall mission package, and the Aircraft Commander or Flight Lead who will lead the attack. In simulation-based CPXs, technical controllers are employed to fill in these data fields. Further, different simulations require distinct fields and data formats –so the technical controllers must know a variety of aircraft parameters, Numbered Air Force (NAF) procedures, tactics and the requirements several simulation systems.

The IMCN program is creating a set of knowledge bases to fill in the data currently supplied by controllers. A WOC knowledge base allocates resources (i.e., assign squadrons, aircraft, and weapons), a Package/Mission Commander knowledge base checks timing parameters (e.g., on station, time over target, etc.) and create contingency plans, and an Aircraft Commander/Flight Lead knowledge base engages in detailed mission planning and conducts sanity checks (e.g., altitude assignments, refueling offloads, etc.).

In addition, the IMCN program is developing a high level architecture (HLA) command and control data interchange format (C2DIF). The C2DIF is a semantic structure capable of providing data to the heterogeneous simulations used by the Air Force, both now and for the foreseeable future. The C2DIF was created after reviewing data requirements of existing and proposed simulations that flight-plan or launch air missions. The data fields of the ATO and ACO required from the Air Operations Center (AOC) were also analyzed then combined with the simulation requirements into a single analysis sheet. ATO and ACO data fields that had no bearing on the control of air missions in the targeted simulations were eliminated while data missing from the AOC US message text formatting message was added for each simulation. C2DIF eliminates the need for controllers to learn specific simulation syntax and allows them to concentrate on the planning details needed to provide operational realism in the flight plan.

Finally, an advanced user interface was created to help controllers create rules and alter the IMCN rule base to reflect current exercise goals and procedures or accommodate different operational methods and tactics during training.

UNCLASSIFIED

The IMCN should be fully integrated with the Air Force's Air Warfare Simulation (AWSIM), the Navy's Research, Evaluation, and Systems Analysis (RESA) simulation, and the Navy's JSAT simulation by the end of 2001. It will also be an integral piece of the Air Force's next generation simulation, the National Air and Space Model (NASM), when it is fielded in 2002.

Graphical Agent Development Environment II (GRADE II⁴²)

The GRADE program is a Small Business Innovate Research (SBIR) program developing a toolkit to simplify the construction, validation and visualization of the Situation Awareness Model for Pilot-in-the-Loop Evaluation (SAMPLE) based human behavior models.

The GRADE effort is expanding and enhancing the SAMPLE agent architecture to provide additional cognitive modeling capabilities and tools to ease model construction and maintenance. It will provide the modeler with full access to all levels of the SAMPLE architecture in an intuitive graphical manner. In addition, a specific agent model relevant to the USAF Test and Evaluation (T&E) community will be developed and integrated within a USAF-sponsored simulation environment to provide for a full demonstration and evaluation of the combined SAMPLE/GRADE software.

Enhancing the Usability of Computer Generated Forces⁴³

Aptima is creating CGF nodes that can function effectively as team players in a Distributed Mission Training (DMT) environment. The CGF models are being designed to support DMT team-training objectives by working as "confederates" with scenario designers and exercise controllers. The models will be developed using BBN Technology's D-OMAR (Distributed Operator Model Architecture) framework, and tested in experiments using Aptima's Distributed Dynamic Decision-making (DDD) testbed that represents an AWACS (Airborne Warning and Control System) team task. This is a 2-year Small Business Innovative Research (SBIR) project.

UNCLASSIFIED

One of the key goals is to identify target DMT training objectives by reviewing Aircrew Mission Essential Competencies definition documents. A DDD Team-in-the Loop Simulation test environment will then be created as the command and control environment using CGF models created with the D-OMAR system. Individual models will be created that embody different instructional strategies and their training effectiveness will be tested by comparing performance of subjects trained with human teammates to performance of those trained by a simulated confederate. Specific test goals include: assessing whether the human participants exhibit the desirable behaviors when they were trained with the models; assessing whether the “hybrid” teams of human players and models achieve levels of team performance comparable to those achieved by all-human teams; assessing whether the behavior of the models is credible to the live players; assessing the consequences of different instructional strategies. The product of the effort will be a series of team-player models for a variety of aircrew team roles and a roadmap for applying both the approach and the results in high-fidelity DMT environments.

United Kingdom

Human Factors

Capture of Realistic Behaviors for inclusion within Synthetic Environments (SE)

This program is exploring a number of issues related to measures and measurement of the realism of CGF behavior. The task seeks to provide justification for the belief, articulated in the strategy for the development of the MOD synthetic Environment Program (Belyavin and Farmer, 2000) that “...the realistic representation of

⁴² Principal Investigator: Lt John Camp, AFRL/HESS

UNCLASSIFIED

human behavior is a key element in the future development of synthetic entities for defense applications". The task colloquially re-christened "Credibility of Synthetic Forces (CSF)" reflects the central role that the "realism" of CGF behavior is believed to play in stakeholder perceptions of SE credibility and the effectiveness of SE applications.

Reference

Belyavin, A.J. & Farmer, E.W. (April 2000) A strategy for the development of TG-5: human sciences and synthetic environments (FY00 up-date).
DERA/CHS/PPD/CR000282/1.0.

Supporting Effective Communication and Interpretation of Command Intent

This research will review the mechanisms underpinning the key characteristics of Command Intent, its communication and interpretation, and those individual or external factors that affect delivery, leading to the development of a theoretically based causal model. This model will be built on data from a literature review, a questionnaire and an interview study, followed by a series of man-in-the-loop micro-studies. The model will identify how commanders prepare intent statements; what information typifies effective intent statements; and conversely, what the consequences are should that information be absent. Existing military policies and procedures will be examined to identify how they could be enhanced to further support the effective communication and interpretation of command intent. In addition, training methods across all three services will be reviewed for information concerning best practice in relation to the teaching of command intent. Procedural, training and technology design recommendations will be made based on the outcome of these studies.

Team Processes

This program of research will provide clients with structured tools and methods to help design appropriate teams to accomplish tasks in specified military domains. The tools will provide a way to characterise and quantify the impact of alternative team

⁴³ Principal Investigator: Jean MacMillan, Aptima, Inc.

designs and structures on combat effectiveness as well as the consequences of team design on team processes.

The work will proceed in two parallel streams: development of a theoretical model of team structure and associated optimisation procedure; and validation of the theoretical model using Integrated Performance Modelling Environment (IPME). A simple model of generic teamwork constraints will be developed including those that can be measured relatively easily, for example communication, and a means devised to 'tune' team structures for optimal performance. In parallel with the development of the theoretical model, recent work on the Simulation of Team Mental Models (STEMM) using IPME will be examined to determine its applicability to this work. Ongoing work examining team performance shaping factors will also be reviewed for potential application to the study.

Doctrine

Manoeuvre doctrine has four driving principles: tempo, simultaneity, firepower, and surprise. Of these, surprise – having the largest psychological component - has the greatest potential for high pay-off from low investment. An academic study of the empirical bases of surprise will be conducted to quantify its effect, such as the divergence of expectation from occurrence and subsequent physiological or behavioural responses. This will be supported by a focused historical analysis study and observations of training exercises. This will be followed by three phases of work. The first will use the core of the conceptual model, and any empirical evidence from the academic study, to begin shaping a mathematical representation of surprise in IPME. The second phase will involve data collection from field exercises in order to validate the model. The third phase will use the results of the validation and data gathering to complete the shaping of the mathematical model and populate it with further empirical data. The final task will be to include the validated model into a variety of OA (operations analysis) tools.

UNCLASSIFIED

In FY2001/2002 the UK Corporate Research Program is intending to address the issues of improving the representation of human variability within CGFs, and identifying the UK MOD requirement for SE-embedded CGFs.

Canada

Development of HBR models in CA is usually a component of a larger project and these models are built to address specific needs of those projects. No current projects directed towards development of HBR models for general application were reported within CA (Defence Research and Development Canada, the Operations Research community or the Canadian Forces experimentation units) although such programs are being considered. Currently, no integrated approach to M&S exists in CA for the CF⁴⁴ although the Canadian Forces Experimentation Centre is being established with a mandate to support and coordinate joint experimentation for the CF and HBR is expected to be a major component of the CFEC effort. There are human behavioural research studies (similar to the US 6.1/6.2 R&D classification) underway that could be used to develop and validate HBR models. Many of the following projects are just beginning, with little published work. The designation following each project title (e.g., 12kr immediately below) is an index code for use when requesting project-related documents.

Army Command and Control Information Systems (CCIS) Development and Training Through Simulation (12kr)

This project will demonstrate an HLA simulation federation to study distributed simulation processes, combining command agents and human participants in simulations with an emphasis on operations other than war. A model of the decision making process in military command will be developed for units below the brigade level (battalion, company, platoon and section levels) using contemporary AI approaches. The command

model will include both allied and enemy doctrine as well as non-military entities such as insurgents, refugees, organized crime, non-government organizations, and the like. The project will exercise the models in a distributed intelligence simulation involving a live field exercise as well as experiment with various levels of HLA compliance and 3D visualization techniques. Supporting work exploring the use of command agents based on naturalistic decision-making is underway at the Royal Military College in Kingston.

References

Liang, Y., Robichaud, R., Fugère, B.J., & Ackles, K.N. (2001) Implementing a naturalistic command agent design. *Proceedings of the 10th Conference on Computer Generated Forces and Behavioral Representation*, pp. 379-386, Norfolk, VA.

Tactical Aviation Mission System Simulation (TAMSS) TD (13cp)

The objective of this technology demonstration is to link high-fidelity, constructive models through a distributed, interactive M&S environment with crew-in-the-loop helicopter simulators to support acquisition decisions, investigation of system integration issues, rapid prototyping to resolve user-interface design issues and the development or modification of doctrine to ensure that the effectiveness of these new technologies is maximized. TAMSS will incorporate a Wide Area Network communication infrastructure to facilitate a wide range of evaluation domains, including human-in-the-loop operation, multi-ship or multi-site scenario execution, equipment or hardware-in-the-loop operation, or any combination thereof. Human Behaviour Modelling within the TAMSS program will be done through entity behaviour scripts within the STAGE™ application (<http://www.cae.com>).

⁴⁴ Creating the Canadian Forces of 2020 symposium Modelling and Simulation: Enabling the Creation of Affordable, Effective 2020 Canadian Forces, Section 4, US, AS, UK, FR, NL, and NATO strategic M&S approaches, http://www.vcds.dnd.ca/dgsp/dda/symp/ms/sec4_e.asp

UNCLASSIFIED

Maritime Air Littoral Operations (MALO) Technical Demonstrator (13d)

The objective of this project is to demonstrate the use of emerging distributed M&S technologies and processes. The demonstration focuses on modeling, development and assessment of the Maritime Air Component of maritime littoral operations in context of the total task force and environment tactics and doctrine. Specific studies will involve developing an HLA environment to study preliminary concept of operations for new capabilities being provided by Maritime Helicopter Project (MHP) and the Aurora Incremental Modernization Project (AIMP) against land, surface and sub-surface threats. MALO will develop models of specific mission segments to sufficient level of fidelity so that analysts may experiment with different human performance models to study the effect of various assumptions and modeling approaches on predicted performance. Specific studies planned will be an ASW (Anti-submarine Warfare) screening task, examining concept of operations and tactics for the use of multi-static acoustics, and a coastal over land surveillance task, examining the optimal use of imaging radar and possibly the fusion with other sensors. Crew performance and decision making issues will be included in the models as part of the concept of operations development and tactics assessment as well as the performance of task force C2 structure. Human behaviour will be represented mainly at the aggregate levels of crew concept as opposed to individual operator level at a work station, however for certain sensor functions and tasks, individual operators may be modelled. Links to physical models and the capability to introduce limited man-in-the-loop participation will be included to study on board crew tasks and interactions

Cockpit Systems Integration (13ia)

The objective of this project is provide a technology-base in cockpit systems integration through human modelling and simulation to support acquisition, technology of advanced flight simulation and airborne simulation research. This project develops tools, methods and procedures that support the procurement, update, or modification processes in the life cycle management of CF airborne platforms, incorporating human factors through Simulation Based Acquisition (SBA). Human behaviour is modelled using IPME (www.maad.com) and data is collected through a companion tool, the

UNCLASSIFIED

Human Factors Analysis Tool (HFAT), that is being developed to support HFE (human factors engineering) analyses comparable to that outlined in Mil HDBK 46855A.

Verification and validation of the IPME environment and its embedded models are in progress. The current effort is examining the validity (Cain & Hendy, 1998) of IPME's IP/PCT (Information Processing/Perceptual Control Theory) workload model (Hendy & Farrell, 1997) against empirical data from UK psychology experiments that were used to develop the UK POP (Prediction of Operator Performance) workload model.

References

Cain, B. & Hendy, K. C. (1998). *A protocol for validating task network simulation procedures* (TTCP/HUM/98/0004): The Technical Cooperation Program, Group HUM, Technical Panel 7.

Hendy, K. C., Farrell, P. S. E. (1997). *Implementing a model of human information processing in a task network simulation environment* (DCIEM No. 97-R-71). Toronto: Defence and Civil Institute of Environmental Medicine.

Mil Handbook 46855A (17May1999) *Department of Defense Handbook, Human Engineering Program Process and Procedures*. Commander, US Army Aviation and Missile Command, ATTN: AMSAM-RD-SE-TD-ST, Redstone Arsenal, AL 35898-5270. <http://astimage.daps.dla.mil/online/>; <http://astimage.daps.dla.mil/docimages/0001/83/06/46855A.PD6>

Advanced distributed mission training (ADMT) technical demonstrator (TD) (16be)

The ADMT TD will advance the CF knowledge of simulation for training in a distributed mission-training environment and begin to redress some of the shortfalls in air-combat training capabilities for CF18A aircrew. The ADMT will make use of new, cost-effective, distributed M&S systems, supporting both research and operational requirements in three key areas: high-fidelity visual displays, intelligent agents and distributed simulations. Collaborative work between CA and US is underway to develop a CF18A ADMT cockpit simulator as well as a preliminary assessment of intelligent

UNCLASSIFIED

agent tools that might be used to provide both red and blue virtual forces to support research of synthetic training issues.

Representation of human behaviour (16bx)

This project will develop cognitive and physical behavioural models of individuals or groups in various constructive simulations. Part of this project will provide support to the ADMT technical demonstrator (16be), however, the mandate is broader than fast-air. Initially, work will concentrate on developing a virtual helicopter deck-landing simulation (HDLS) environment comprising human-in-the-loop and virtual agents. When the ADMT CF18A simulator becomes operational, lessons-learned in the HDLS project will be used in the development of complementary capabilities in the ADMT. One of the aims of this project is to explore the use of constructive agents in training applications to substitute for instructors or ancillary crew.

Command-team skill, decision making and leadership (16ka)

This project seeks to identify the critical aspects of human cognitive behaviours in the context of new control technology, changing operational environments and evolving military doctrine. The effect of fatigue on teamwork has been studied as well as two studies of confidence measures in threat assessment. The project will explore factors such as confidence and expertise, creativity in a team environment and decision-making under uncertainty, contributing to the current body of knowledge on Command and Control to establish a model of how these factors affect team performance. The effect of sleep-deprivation on both individual and team performance in judgement and decision-making tasks is being studied to expand our knowledge of teamwork in a Command and Control environment.

References

Baranski, J. V., Cian, C., Esquive, D., Raphel, C., (1996) *Effets du modafinil sur les performances cognitive au cours d'une privation de sommeil de 60 hours.*

UNCLASSIFIED

Technical Report #7, Centre de Recherches du Service de Santé des Armées (CRSSA), La Tronche. 20 pages.

Baranski, J. V. (1995). Evidence for a Modafinil induced overconfidence effect in sustained operations. *Proceedings of the 37th Conference of the International Military Testing Association*, pp. 115-120.

Pigeau, R., McCann, C., Baranski, J. V., Thompson, M., Taylor, M., Shek, Y., Buguet, A. (1995). Modafinil, amphetamine and placebo as alerting substances for sustained operations. *Proceedings of the 37th Conference of the International Military Testing Association*, pp. 85-90.

Thompson, M., Baranski, J. V. (1995). The effects of modafinil, amphetamine and placebo on individual and group decision-making processes in a sustained operations environment. *Proceedings of the 37th Conference of the International Military Testing Association*, pp. 103-108.

Baranski, J. V. (1997). *Information order effects in TANDEM II*. In R. Pigeau (Ed.) *Proceedings of the TTCP- UTP5 workshop on Factors Affecting Human Information Processing*, Toronto, Canada.

Raphel, C., Esquivie, D., Stivalet, P., Cian, C., Baranski, J. V., Leiffen, D., Poquin, D., Barraud, P. (1998). *Modafinil effects on spatial cognition during 60 hours of sleep deprivation*. *Proceedings of the Advisory Group for Aerospace Research and Development (AGARD)*, CP-599: Aeromedical support issues in contingency operations. Rotterdam, The Netherlands (10.1 – 10.4).

Cian, C., Raphel, C., Baranski, J. V., Esquivie, D., Stivalet, P. (1997). Sustained military operations: Psychological effects of modafinil during sleep deprivation. *Proceedings of the 38th Conference of the International Military Testing Association*, San Antonio, Texas.

UNCLASSIFIED

Baranski, J. V., Petrusic, W.M., (1994) The calibration and resolution of confidence in perceptual judgments. *Perception and Psychophysics, 55*, 412-428.

Baranski, J. V., Pigeau, R., Angus, R., (1994). On the ability to self-monitor cognitive performance during sleep deprivation: A calibration study. *Journal of Sleep Research, 3*, 36-44.

Baranski, J. V., Petrusic, W.M., (1995) On the calibration of knowledge and perception. *Canadian Journal of Experimental Psychology, 49*, 397-407.

Pigeau, R., Naitoh, P., Buguet, A., McCann, C., Baranski, J., Taylor, M., Thompson, M., Mack, I., (1995) Modafinil, d-amphetamine, and placebo during 64 hours of sustained mental work: I: Effects on mood, fatigue, cognitive performance and body temperature. *Journal of Sleep Research, 4*, 212-228.

Baranski, J. V., Pigeau, R., (1997) Self-monitoring cognitive performance during sleep deprivation: Effects of Modafinil, d-amphetamine, and placebo. *Journal of Sleep Research, 6*, 84-91.

Petrusic, W. M., Baranski, J. V.(1997). Context, feedback, and the calibration and resolution of confidence in perceptual judgments. *American Journal of Psychology, 110*, 1997, 543-572.

Baranski, J. V., Cian, C., Esquive, D., Pigeau, R., Raphel, C., (1998) Modafinil during 64 hours of sleep deprivation: Dose related effects on fatigue, alertness, and cognitive performance. *Military Psychology, 10*, 173-193.

Baranski, J. V., Petrusic, W. M., (1998) Probing the locus of confidence judgments: Experiments on the time to determine confidence. *Journal of Experimental Psychology: Human Perception and Performance, 24*, 929-945.

UNCLASSIFIED

Baranski, J. V., Petrusic, W. M., (1999). Realism of confidence in sensory discrimination. *Perception and Psychophysics, 61*, 1369-1383.

McCann, C., Baranski, J. V., Thompson, M. M., Pigeau, R., (2000). On the utility of experiential cross-training for team decision making under time stress. *Ergonomics, 43*, 1095-1110.

Baranski, J. V., Gil, V., Moroz, D., McLellan, T. (2000). Effects of modafinil on cognitive performance during 40 hours of sleep deprivation in a warm environment (pp. 113-136). In A. Buguet, M. Radomski, & D. Moroz (Eds.) *Effects of modafinil on thermoregulation during physical exercise and sleep deprivation in a hot environment*. DCIEM: Toronto, Canada (ISBN: 0-9688317-0-2).

Stress and coping in CF personnel during operations (16kb)

The objective of this project is to develop and validate a conceptual model of personnel response to operational stresses and to develop countermeasures to those stresses. A model of stress and coping has been proposed and is presently being validated. Data collection on stress and performance is ongoing to support the development of individual styles of coping with stress.

References

Thompson, M.M., Gignac,M.A.M. (2001). Adaptation to Peace Support Operations: The Experience of Canadian Forces Augmentees in P.J.M.D. Essens, A.L.W. Vogelaar, E. Tanercan, D.J. Winslow (Eds.).*The Human in Command: Peace Support Operations*. Amsterdam: Mets & Schilt.

Thompson, M.M., Gignac,M.A.M (2001). *A model of psychological adaptation in Peace Support Operations: An overview*. DCIEM Technical Report (TR 2001-050). Toronto, Ontario: Defence and Civil Institute of Environmental Medicine.

Human Engineering performance modelling environment (16kc)

This project's goal is to develop computer-based models and tools to apply human factors knowledge to system design and acquisition. A major focus of this work will be the inclusion of team-relevant factors and tasks in a task network simulation environment (IPME) so that models of team performance can be studied together with models of individual performance. The Netherlands Organization for Applied Scientific Research (TNO)⁴⁵ and DCIEM (Defence and Civil Institute for Environmental Medicine) have begun collaborating on team modelling, with TNO developing a link model of team factors and activities using a naval frigate as a case study.

Organizational and social issues (16kx)

The objective is to develop a theoretical framework for Command and Control to guide training programs. The proposed approach will investigate factors such as interpersonal trust and the sharing of intent as they affect command and the effectiveness of CF command teams. This may provide a defensible approach to the introduction of objective models of command factors in CGF. This project is only planned and is not yet funded although some preliminary work is underway.

STRIVE (<http://www.cae.com>)

STRIVE™ is an HLA-compliant, tactical battlefield simulation software package intended for military training and analysis, developed and marketed by CAE Inc⁴⁶. STRIVE™ is PC/Windows NT based, integrating participating federates into common exercises and controlling the behaviours of computer generated forces (CGF). The synthetic environment can represent air, sea or land, however, the template environment and sample CGF focus on army and army-aviation operations. STRIVE™ is intended as a

⁴⁵ http://www.tno.nl/homepage_nl.html

⁴⁶ CAE Electronics Ltd., C.P. 1800 Saint-Laurent, Quebec, Canada, H4L 4X4 <http://www.cae.com>

UNCLASSIFIED

future replacement for ITEMS™, providing all of the ITEMS™ capabilities as well as additional features and improved usability.

CGF behaviours are controlled through its “doctrine”, a modular, rule-based set of instructions that are built through the Strive-Studio. Additionally, an “Expert System” framework has been provided to integrate existing cognitive behavioural models. Several doctrine models may be created with differing rule sets, and then a doctrine can be associated with an entity during the scenario creation. Any combination of platform, sensors, weapons systems, countermeasures, communications and default behaviours can be combined to create a player. Inter-player, digital communications can be modelled, including tactical messages that comprise information or instructions to other entities. The rule consequences are organised into three categories: navigation, communication and weapon firing. A player can execute one consequence or action from each category at a time and the rules are executed in order of presentation with the categories. Although a rule set can invoke another rule set, it can only be a subsequent rule in the hierarchy.

STRIVE™ is relatively new, and has few native libraries of CGF behaviours. Defence Research and Development Canada has selected it for beta testing in current SBA and training applications.

IPME – Integrated Performance Modelling Environment (<http://www.maad.com>)

IPME is a discrete-event, network simulation software being developed collaboratively among the United Kingdom's Defence Evaluation Research Agency's Centre for Human Sciences (DERA CHS, now QinetiQ), Canada's Defence and Civil Institute for Environmental Medicine (DCIEM) and Micro Analysis and Design Inc. (MAAD) from Boulder, CO.

The IPME uses a process-oriented modelling approach and builds upon an SME's accounting of how operator activities are organized to meet operational objectives. Operator responsibilities and goals can be recorded at a high level of abstraction (such as "Prepare for Mission") that can be decomposed into a hierarchy of functional blocks

UNCLASSIFIED

(such as "Prepare Met Brief") until the analyst has reached a level of granularity (such as "Read Current Weather Map") appropriate to study a given problem.

The analyst can incorporate HBR sub-models to dynamically modify individual operator task "Time to Perform" and "Probability of Failure" values. Currently, IPME only provides the analyst with the capability to enter and use HBR sub-models; no performance shaping factor behavioural model is supplied. IPME includes micro models of basic human actions that can dynamically set the time required to perform a task.

IPME incorporates two models of operator workload: DERA CHS's Prediction of Operator Performance (POP) (Jordan et al., 1996) and DCIEM's Information Processing (IP) model (Hendy & Farrell, 1997). Both models use similar concepts of Time Pressure resulting from task demands to represent workload, although the approaches differ. Each approach constrains task execution, scheduling tasks to better mimic operator response by delaying, rescheduling and shedding tasks as concurrent-task demands exceed operator capabilities and operator time pressure approaches 100%.

The POP model is based on and validated through empirical studies (primarily experimental psychology studies supplemented with simulator studies), using the results to formulate rules for combining subjective operator ratings of Input, Central and Output task demands. The combination rules are invoked through a Markov model to capture task-interference effects, combining demands in the Input, Central and Output phases of pairs of active operator tasks to predict a composite perceived workload in each phase.

The IP/PCT model is based on commonly accepted human factors and psychology literature although it has yet to be validated. The time to complete a task is prolonged when concurrent tasks interfere by drawing on similar mental resources. Tasks are prioritized and a maximum duration is assigned; time pressure is calculated by the time required to complete a task divided by the time available and the operator's instantaneous time pressure is the maximum time pressure of all active tasks.

References

Hendy, K. C., & Farrell, P. S. E. (1997). *Implementing a model of human information processing in a task network simulation environment* (DCIEM No. 97-R-71). Toronto: Defence and Civil Institute of Environmental Medicine.

Jordan, C. S., Farmer, E. W., Belyavin, A. J., Selcon, S. J., Bunting, A. J., Shanks, C. R., & Newman, R. (1996). Empirical validation of the prediction of operator performance (POP) model. *Proceedings of the 40th Annual Meeting of the Human Factors and Ergonomics Society*. Santa Monica, CA: Human Factors Society.

Australia

Joint Simulation Environment

ADSO plans to sponsor the development of a Joint Synthetic Environment Concept Demonstrator, involving both government and industrial partners. This will include the linking of DSTO's simulation research facilities, such as Land Operations Division's Synthetic Environment Research facility in Adelaide and Air Operations Division's Melbourne-based Air Operations Simulation Centre. The Joint Simulation Environment will be a multi-service focussed environment where simulation can be conducted serving the broader ADF (Australian Defence Force) simulation community.

Architecture Studies

Defence envisages an increased integration of simulation that is likely to require the combination of discrete simulations via physical interconnection across computer networks. Connection to the simulations of allied nations is also envisaged. Distributed Interactive Simulation (DIS) and High Level Architecture (HLA) both provide the ability to achieve this interconnection. While most ADF training simulators currently plan to

employ DIS to provide networking capability, migration to the newer HLA standard may be required to achieve interoperability with simulators from Allied nations. Thus DIS/HLA migration is critical to most training simulators in the Australian Defence Department.

Simulation Standards Studies

The selection and implementation of suitable standards is one way in which Defence can improve its management of simulation. Effective standards can reduce the cost of simulation ownership, by reducing duplicative effort and enhancing re-use of simulation elements. Standards also facilitate the cost-effective inter-operability of simulations. Even if simulations are not directly connected via networks, their effectiveness may be enhanced if they comply with a common set of standards.

There are a number of areas where simulation standards exist, such as in the areas of simulation interconnection and visual system databases. Competing, disparate or complementary standards do exist. Conversely, in some areas of simulation there are various formats or solutions which are used but which are not identified or recognised as being a ‘standard’.

Consequently, a conscious and methodical study is required to determine, against the backdrop of current and future simulation activity in Defence, a list of simulation standards to be adopted by Defence.

Intelligent Agents Studies

The technology of intelligent agents appears to have considerable merit for the development of Joint Synthetic Environments. There is a current proposal by Industry to ADSO (Australian Defense Sciences Organization) to investigate the feasibility of using intelligent agents to manage distributed simulation environments, and there are current industry contracts within DSTO seeking to enhance the use of intelligent agents within simulations. The intent of this work is to allow more autonomous unit behaviours within

wargame environments, and eventually to permit the same agents to be applied to close simulations. The work holds great promise to reduce the number of military staff required to host an experiment, as well as introducing greater flexibility and fidelity to the automatic behaviours.

Virtual Air Environment

The Virtual Air Environment (VAE) concept integrates real assets and virtual simulations (comprising Human-in-the-Loop (HIL) and computer-generated entities) in one synthetic environment to create a virtual world for training and other purposes. A number of new capabilities are due to enter RAAF service in the next few years, including the Airborne Early Warning and Control (AEW&C) and Jindalee Operational Radar Network (JORN). VAE will make extensive use of modelling and simulation to provide a training system that will overcome many of the projected shortcomings of the current training systems by providing a synthetic environment in which live assets can be stimulated by, and interact with, networked simulators or CGFs. To determine the viability of this concept, an Initial Development Phase (IDP) is under-way to investigate the technical and operational issues associated with this proposal. There are two aspects to this work:

- linking distributed HIL simulators with operational systems
- stimulating the operational system with computer generated entities from a constructive simulation or other scenario generator

The feasibility of using CGFs to provide training in the VAE is being explored through the development of prototypes in the simulation environment BattleModel. It is considered that Battle Model and the use of intelligent agents could provide very sophisticated and tactically credible CGFs for the VAE.

BattleModel has a DIS interface to enable its entities to interact with other VAE entities. Intelligent agents, developed using dMARS during AEW&C work (described below), have been adapted to control the behaviours of the CGFs.. High-level control of

UNCLASSIFIED

the Blue Force CGFs is also possible through a GUI that sends commands to the CGFs Agents. Orange Force CGFs (strikers and their escorts) behave autonomously in the current scenarios.

A JAVA GUI for controlling Blue Force CGFs through the BattleModel radio communications model has been developed that allows operators to issue simple commands to the CGFs such as “scramble”, “scramble to CAP”, “vector”, “return to base”, “change speed” and “change altitude”. This interface provides the minimal level of control required by trainee air defence controllers to practice basic air intercept geometries in the VAE.

The intelligent agents controlling the BattleModel entities respond to commands but also provide a high degree of autonomous behaviour. The agents are used to model pilots of the various simulated entities (fighter and strike aircraft for friendly and hostile forces) and they respond appropriately to commands issued through the GUI described above. Their autonomous behaviour includes flying waypoints, flying in formation, engaging opposing force aircraft, and firing missiles in an attempt to destroy opposition forces.

Figure 2 shows the first intercept by Blue fighters (call sign Shogun) against the Orange strikers. Missiles launched by Blue fighters appear as unknowns with yellow symbology and have names derived from the name of the launch platform. At this stage, the second pair of Blue fighters (Viper) is on CAP (Combat Air Patrol) to the North West of Tindal.

UNCLASSIFIED

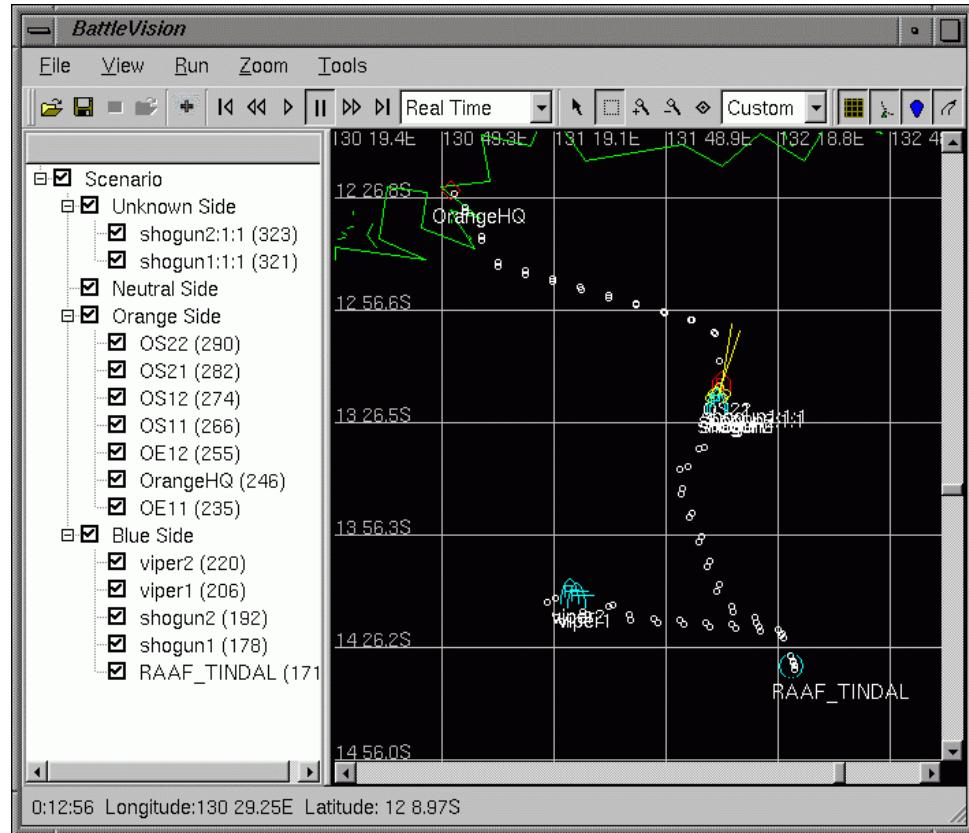


Figure 2. Intercept of Northern strikers.

Figure 3 shows the first engagement of the escorted pair of Orange strikers by Blue Fighters (Viper). This Orange strike package is tasked to attack the Blue base at Tindal. They are approaching from the west and heading directly east towards their target. The Orange escorts have diverted from their flight path to engage the Blue interceptors.

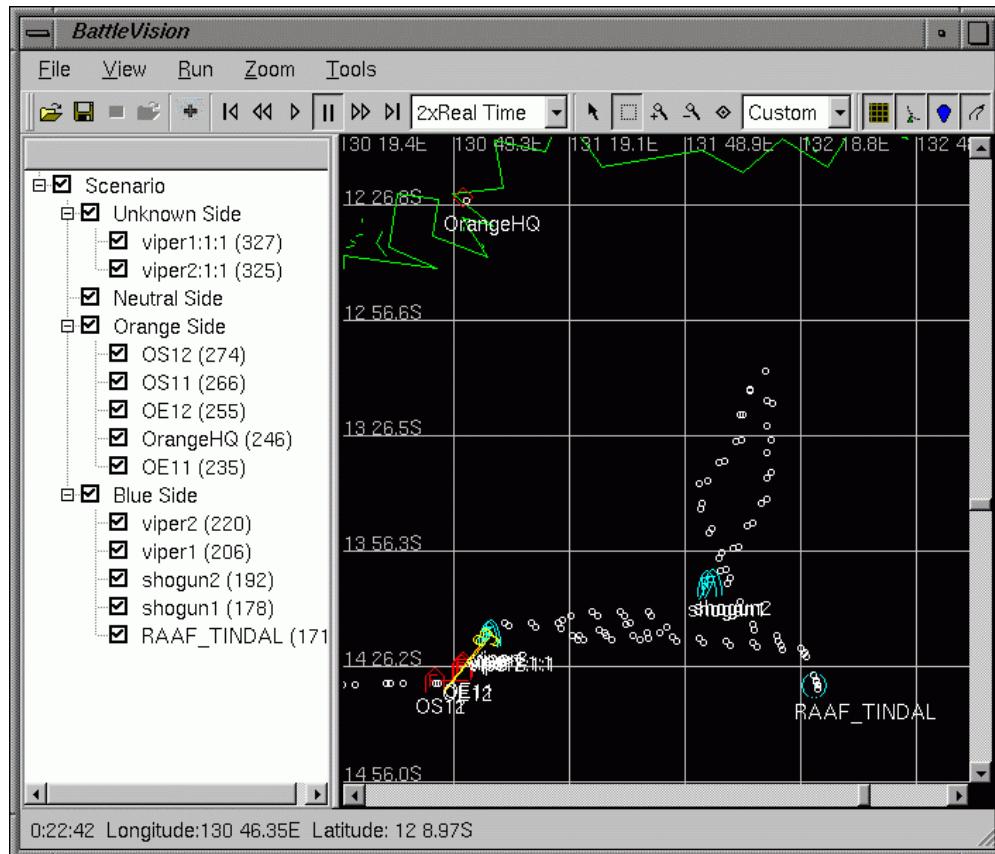


Figure 3. First intercept of escorted Orange strikers during demonstration.

The AEW&C Agents used for the VAE were modified to accept commands through the JAVA GUI. These Agents have scaled well to the VAE scenarios and HiL environment. Scenarios for a fully developed VAE will be on a much larger scale with many more entities involved. The CGFs required for these scenarios may be required to exhibit potentially complex and tactically credible behaviours to ensure humans trainees receive appropriate, accurate responses. Intelligent Agents are likely to be a key technology in achieving this goal but how well the current Agent technology will scale to a fully developed VAE is yet to be established. There are also likely to be HiL simulation issues that are not typically encountered in the OR domain such as a requirement for agents to interact with the simulated environment (ie. terrain, topographic features, etc) to

provide appropriate visual displays for participants. This may require new approaches so the Agents can perceive the simulated environment.

AEW&C (Project Wedgetail) Acquisition

Agent based simulation has been used for a long time in operational analysis studies supporting the Airborne Early Warning and Control (AEW&C) project. The emphasis of these studies was initially on the evaluation of comparative hardware solutions, requiring only simple agents with shallow command and control structures. These first agents merely had to operate the hardware in a reasonable manner consistent with the system capabilities. Simple procedural languages would have been adequate but a richer repertoire of behavior was anticipated for future studies. The emphasis of these studies turns to the development of operational concepts and tactical procedures as the hardware definition is refined, requiring much more realistic representation of both individual and team behaviour.

Constructing an agent-based representation of an AEW&C team is quite a challenge and has not yet been fully attempted. The early agents were only required to perform sub-tasks, sufficient to achieve proper operation of the equipment with formalised, simple relationships between agents. We are now moving to multi-agent systems that have depth in both command and control structure as well as in time.

The depth of command and control structure requires that agents be capable of packaging and delegating tasks to those agents that have the capability to understand the requirements, read the situation, and carry out the intention of the delegating agent. These delegating agents must in turn be capable of understanding the intention of the agent above them in the command control structure in the tasks delegated to them.

By "depth in time" we mean the ability of an agent to progressively gather an understanding of the situation over a period of time, infer the probable intention of those forces are under direct control and project the situation forward in time.

Maritime Surveillance

Air Operations Division (AOD) is supporting the RAAF Maritime Patrol Group in developing new tactics and concept of operations for the upgraded AP-3C Orion Maritime Patrol Aircraft. The Orions are used by the RAAF in peacetime for maritime search, surveillance and intelligence gathering operations in and around Australian territorial waters including the South East Asia and Pacific regions.

The Orions are currently undergoing an extensive upgrade program that includes new sensors and avionics that have the potential to significantly improve the capability of the aircraft. Because Maritime Patrol Groups have no previous operational experience with some of these new sensors, AOD is working closely with them to baseline the expected mission performance of the aircraft in typical mission profiles and scenarios, and also to develop new, integrated flying and sensor employment policies that allow the aircraft to function at its full potential.

The requirement from the RAAF was for AOD to investigate the effectiveness of different flying tactics and sensor employment policies as a function of weather conditions, geography and various other factors that impact on mission effectiveness. In order to satisfy all these requirements it was necessary to have a very detailed physical model of the aircraft and its sensors, including realistic flying characteristics, fuel consumption and sensor performance. It was also necessary to have a model of the tactical decision making process on board the aircraft that was capable of representing the actual human operator and crew workload (including the type and amount of information available at any given time), the sensor data-fusion process and chain of command. The tactical decision making model also had to be flexible and robust enough to allow timely and easy modification to investigate different operational procedures and tactics.

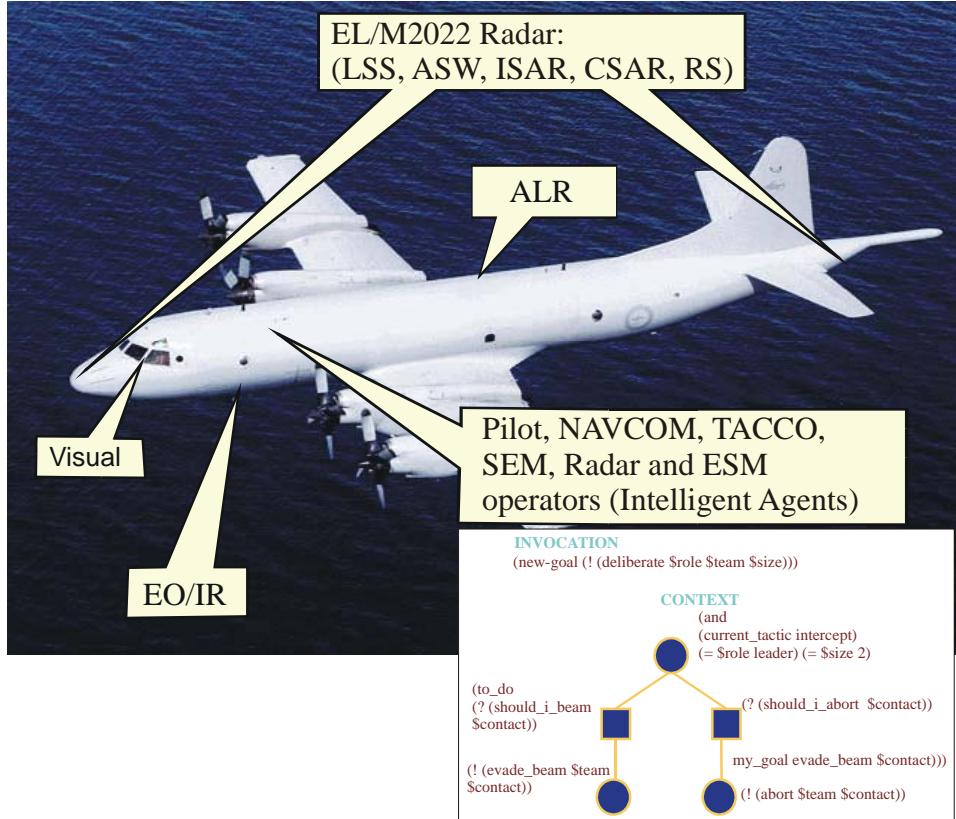


Figure 4. The various sensors and human operators that are modelled in this work superimposed on a photograph of a P-3C Orion maritime patrol aircraft.

The BattleModel simulation framework was chosen as the primary modelling environment for this work, with the tactical decision making model constructed using individual Intelligent Agents (implemented in the dMARS language) for six crewmembers with significant roles in the tactical decision making process; communication between the agents was modelled via a radio intercom link.

Intelligent Agents were chosen because of the requirement to model decision-making based on a degree of awareness of the environment or tactical situation the aircraft finds itself in. Maritime surveillance tactics, as with almost all tactical decision making, rely on making an assessment of the current situation based on fusing data from different sensors and also on making some assessment about the intent of other entities in

the environment. The Beliefs, Desires, Intention formulation of the dMARS agents readily lends itself to this type of modelling..



Figure 5. Six RAAF P-3 crew members from Maritime Patrol Group forty-six minutes into a simulated mission using the ‘Battle Model’ in interactive ‘crew-in-loop’ mode to evaluate some maritime search and classification tactics.

Battle Model is used both for faster than real-time constructive Monte Carlo simulations of missions, with the Intelligent Agents making the tactical decisions, as well as interactively or “crew-in-loop”, where tactical decisions are made by actual crew members. The constructive mode is used to gather information on several hundred simulated missions for statistical analysis and robustness tests of various tactics.

The interactive or “crew-in-loop” mode is used to test and evaluate new tactics in a realistic environment and to refine existing tactics based on statistical analysis of constructive simulation results. In this mode, the tactical picture is projected onto a large screen showing the current sensor information (radar, ESM contacts, aircraft location and state, et cetera) superimposed on to a geographical map of the region. This allows the crew to focus on developing and evaluating higher-level tactical procedures rather than on low-level interactions with individual controls (Figure 5).

UNCLASSIFIED

Figure 6 illustrates the tactical development cycle used by the RAAF and where BattleModel fits into the process. Initial tactics, mission profile and measures of effectiveness are supplied to AOD for initial modelling and analysis at the top of the cycle. Constructive Monte Carlo simulations over hundreds of mission are then performed using BattleModel for a given tactic or environmental condition. The outcomes of these missions are then analysed, both statistically and individually, to determine the drivers of mission effectiveness and identify worst-case scenarios. These results are then presented to the RAAF for evaluation and analysis. The initial tactics are then refined further using the ‘crew-in-loop’ mode based on the modelling results and subsequently input into another iteration or round of modelling. Once the tactics are deemed suitably mature, they become operational.

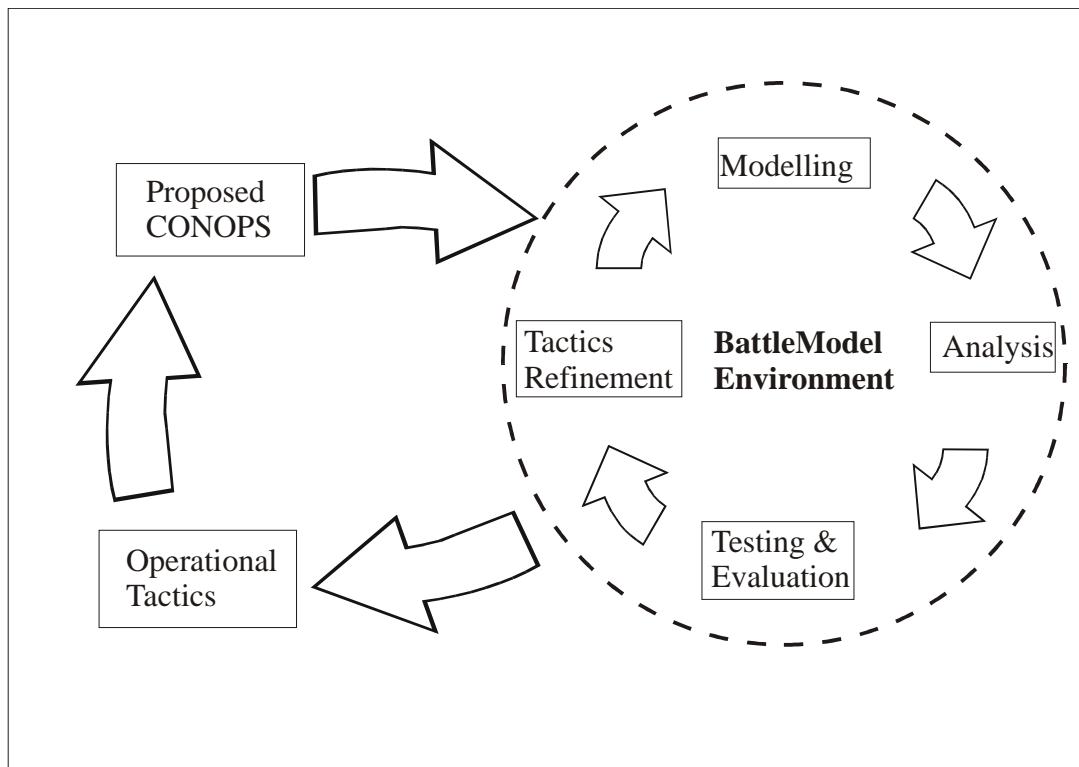


Figure 6. The tactical development cycle used by the RAAF illustrating where the ‘BattleModel’ fits into the process.

Chapter 4: The way ahead in human performance modeling

The first three chapters of this report focus on the need for human performance modeling in the TTCP military communities and summarize current relevant research efforts and capabilities. This chapter focuses on the disparity between identified needs and existing capabilities.

Shortfalls in human performance modeling can be assigned to two broad categories: the behavioral competence and realism of models, and the technologies supporting their application. HBR shortfalls in each category are discussed, noting implications and, where possible, suggesting strategies to address the shortfalls.

Shortfalls in the competence and realism of human performance models

Recent reviews of the current state-of-the-art in human performance modeling for military applications are consistent in the view that present models lack the cognitive and perceptual capabilities and the behavioral realism needed to support cost-effective simulation-based military training and analysis. In their 1998 National Research Council review on representation of human behavior in military simulations, Pew and Mavor point out “the most serious shortfall facing the military Modeling and Simulation community relates not to software, but to the quality and content of the underlying models. Too little attention has been focused on the content of the models themselves, or on the research base needed to create that content”. More specifically, they pointed out that current human models (1) are brittle in that minor deviations from the conditions under which they were created can produce unrealistic behavior, (2) display simplistic responses that do not correspond to the behavior of real people, and (3) lack realistic cognitive abilities. A blue-ribbon panel of military experts in simulation for training, commissioned by the Chief of (US) Naval Research, has drawn a similar general conclusion, emphasizing the generally crude representation of human behaviors in current military simulation

UNCLASSIFIED

“There has been some modeling of textbook doctrine, but much less of how people react to the unexpected. Simulations need better models at every echelon...better models are needed both to simulate opposing forces realistically, and to provide trainees with synthetic colleagues.”
(Wald, 1998)

A more recent NATO report (LTSS SAS-017) extended these conclusions, highlighting the following eight shortfalls related to the modeling of individual combatant behavior

- (1) the inadequacy of current representations of the effects of behavior moderators on human performance(e.g., fatigue, fear)
- (2) the lack of validated empirical data on the effects of moderators on performance
- (3) the absence of mechanisms to infer the intent of opposing forces
- (4) the general lack of integrated cognitive processing architectures (such as situation assessment, decision making, problem solving, or reactive planning)
- (5) the limited capability of models to adapt or learn from experience
- (6) the general absence of self-explanation capabilities—a feature that should improve the training utility of military simulations
- (7) the brittleness of individual models—their inability to respond realistically to unanticipated events
- (8) the predictability of simulated behaviors, created in part by the strategy of modeling “textbook” doctrine rather than combatant behaviors that can be highly variable and only loosely constrained by doctrine

The NATO report also noted that the state-of-the-art in modeling human social structures is even less well developed than that for the modeling of individual combatant behavior. In particular, the report noted that the formal modeling of group, team and organizational

UNCLASSIFIED

modeling is currently in a very preliminary stage of development; further studies are required to characterize the behaviors and interactions of groups, teams and organizations adequately. Particular emphasis was given to (1) the current absence of formal descriptions of team and organization performance moderators and (2) the lack of understanding of the impact of communication or authority structures on the effectiveness of individuals or groups. Without this knowledge, it will be extremely difficult to create valid, generalizable models of the human social structures that are central to the military environment.

These shortcomings are causing training CGFs to be brittle and predictable; their cognitive and perceptual abilities are limited. Training against fully automated simulated adversaries is ineffective because they are simple, predictable and easily defeated training against semi-automated forces (SAFs) requiring many skilled human controllers reduces training flexibility and significantly increases training costs. One example of the latter is a recent large-scale training simulation (Unified Endeavor) that required the use of over 900 human controllers, yielding enormous costs in man-hours and scheduling inflexibility. Recent analyses conducted at the US Naval Air Warfare Center Training Systems Division indicate that the creation of smart, robust, adaptable, unpredictable and challenging computer-generated forces could enable for the first time the capability for realistic anytime, anywhere, on-demand simulation-based training and at the same time achieve greater than a 75 percent reduction in simulation manning requirements.

Several of these human performance modeling shortfalls are discussed further in the following paragraphs, identifying major sources of the problem and suggesting possible approaches to addressing them.

High Level Cognition.

High level cognition refers to the interactive mechanisms of perception and thought in human information processing tasks such as spatial judgement, problem-solving, decision making, situation assessment, planning and inferring the intent of others. Detailed discussions of these and other higher-level cognitive phenomena have

UNCLASSIFIED

been provided by Pew and Mavor (1998) and in the 2001 NATO Technical report on human behavior representation (LTSS SAS-017). A great deal of effort has been expended over the past 50 years to understand and model these separate phenomena. With a few exceptions, however, little effort has been put unto the creation of comprehensive integrated models that extend across the range of human cognitive capabilities.

The notable exceptions are ACT-R, a production-rule based cognitive architecture developed by John Anderson and his students at Carnegie Mellon University over the past two decades, and COGNET, a blackboard-based architecture developed by Wayne Zachary at CHI Systems, Inc. A variety of military tasks have been successfully modelled using COGNET, including shipboard combat information center personnel and the helmsman in a US Marine Corps amphibious assault vehicle (AAAV). ACT-R has less of a history of practical military application, however, adversaries in simulated MOUT (Military Operations in Urbanized Terrain) scenarios are being modelled in an extended version, ACT-R/PM. These examples demonstrate that while relatively little work has been invested in the creation of comprehensive models of cognitive performance, it is now possible. The formula for success in this endeavor, demonstrated most clearly by the COGNET efforts, includes

- (1) the existence of an executable model that can display broad cognitive, perceptual and motor human abilities in operationally realistic contexts
- (2) the domain knowledge base needed to populate the model's knowledge representation system at the level of granularity demanded by the application.

The reasons why so few examples of unified performance models can be cited are, first, very few modellers seriously attempted full spectrum modelling, and second, few military are prepared to investment the time and money to capture the extensive domain knowledge necessary to model military tasks adequately. A reasonable starting point for addressing these problems would be related services within TTCP nations to pool resources supporting a two-phased program. The first phase would conduct detailed cognitive task analyses for a representative set of military tasks, preferably building upon

existing studies. The second phase would create task models, based on the results of the Phase I analysis, in selected modeling environments. This approach could accomplish two things: first, it would provide funds needed to create the domain knowledge required to generate comprehensive cognitive performance models, and second, it would provide a funding incentive to encourage the development and comparative evaluation of comprehensive performance models.

Robustness

The robustness of a human performance model is its ability to respond reasonably to unanticipated events, that is, events not specifically anticipated by the model builder. As noted previously, recent reviews of human performance have highlighted brittleness, the lack of robustness, as an important shortcoming in current performance models. One strategy for achieving robustness is to create a rule base so extensive that all contingencies are anticipated, however, this approach has two problems: First, it is very costly to carry out the extensive knowledge acquisition required to produce comprehensive rule sets for many military tasks; second, it is difficult, perhaps impossible, to anticipate all possible contingencies that might arise in free play military simulations. Two possible approaches that in principle do not share these shortcomings have received attention recently, fuzzy logic and genetic algorithms.

Control systems based on fuzzy rules can provide the same level of robustness as systems of crisp rules up to 100 times larger (McNeill and Freiberger, 1993). Despite this promise, a recent comparison of fuzzy and crisp rule-based human behavior models in the Soar architecture showed little difference in the number of rules for the same level of performance (Laththrop, 2000). It remains unclear why HBR modeling has not benefited from a fuzzy approach although few attempts have been made. Given the promise shown by fuzzy modeling in other domains, we recommend further study through TTCP M&S collaborations.

Genetic or evolutionary algorithms may also assist in producing robust HBR models. These techniques operate on the principle of survival of the fittest, differing

UNCLASSIFIED

primarily in the process “chromosomes” change between generations. They both employ mechanisms such as mutation and crossover of genetic components to create chromosomal variation from one generation to the next. An evaluation value or fitness function is used to determine which chromosomes have the best characteristics; those with high evaluation value tend to survive to reproduce, approaching the characteristics of the fitness function over many generations. A few very preliminary efforts have been undertaken to apply this form of soft computation in evolving human performance models that effectively adapt “chromosomal” rule sets to cover the environment contingencies possible within specified military environments. These evolutionary approaches created adaptive rule sets although the settings, rule sets, and fitness functions used were very limited in scope. It remains unclear whether this approach will improve robustness effectively in more realistically complex military situations. Nonetheless, applications of evolutionary computing approaches to improve HBR robustness merit further research by TTCP nations.

Adaptability

The adaptability of a human performance model is its ability to learn and change behavior based on its experience. Models currently used in military simulations either contain no mechanisms of learning or only a very limited set of such mechanisms. One of the most commonly used human performance architectures used in military simulations, Soar, contains only one learning mechanism—chunking. When Soar encounters a condition for which more than one action might be appropriate, called an impasse, it begins a problem solving process to select the best response. When similar conditions occur, the system takes the same action, skipping the impasse resolution process. Chunking is a very useful mechanism, but few would argue it fully captures the adaptive mechanisms underlying human behavior; additional learning mechanisms are being studied for use in Soar (Laird, Personal Communication). Another rule-based cognitive architecture, ACT-R, has a much more elaborated set of learning mechanisms, including techniques for learning new declarative knowledge, learning from examples, learning by analogy, and the statistical tuning of parameters associated with these symbolic structures

(Anderson, 1993). Modeling of behavior in military settings using ACT-R is just beginning (Lebiere, 2001) so its effectiveness and usability remains to be determined.

Finally, the evolutionary computing techniques described in the preceding section of this chapter are a potentially powerful tool set for implementing adaptability in HBR models. Evolutionary processes could be used to generate the new behaviors to react to unexpected events. Shultz and Greffenstedt (1997) demonstrated the approach using faster-than-real-time processing to produce more appropriate rule sets, yielding robotic systems that effectively adapt in near real-time. Once again, the application of these optimization techniques in human performance modeling is in its infancy, and further study is needed to determine its value to human learning models in military simulations.

Behavior Moderators and Realistic Individual Differences

Very little effort has been put into the creation of performance models that exhibit realistic individual differences in the response of the model to the many variables that moderate human behavior. Pew and Mavor (1998) summarize the issue well:

There are a number of important questions to be addressed about variables that moderate behavior. For example, what is the desirability and feasibility of developing models of humans that vary with respect to personalities, emotions, cultural values, skills, and intellectual capabilities? How do external stressors such as extremes in temperature influence behavior, long hours of continuous work, and exposure to toxins? Can the effects of fear and fatigue on soldier performance reasonably be represented? At the unit level, what are the effects on performance of level of training, the extent to which standard operating procedures are followed, and the degree of coupling between procedures?
(p. 242-243)

These issues represent a shortfall of potential significance for a couple of reasons. First, they tend to produce models that are predictable, compromising training effectiveness. Second, simulated adversaries or team mates all behave in a similar

UNCLASSIFIED

manners, which may not provide the trainee with the range of experiences due to individual differences needed to cope effectively with the full range of human combatant behaviors.

Two classes of performance moderators for HBR models can be distinguished - external and internal. External moderators change model behavior due to demands imposed by the simulation such as physical and mental workload, temperature, sustained noise, sleep deprivation, and toxic substances. Internal moderators include personality, emotions, attitudes, cultural values, intelligence, and level of training. Whereas some empirical and theoretical work has been conducted on both external and internal moderators, the empirical work has been parametrically incomplete, and much of it has been based on data-gathering strategies that are difficult to validate (e.g., retrospective judgments of experienced combatants). The 2001 NATO report on human behavior representation (LTSS SAS-017) nicely summarizes many of the major dimensions along which people can vary as they respond to military conditions.

It is clear that we do not have a good general understanding of the effects of moderators on human behavior. Without this understanding, we cannot decide with confidence if it is desirable to capture these effects in our models. Our knowledge of the impact of key individual moderators on behavior is very uneven, lacking parametric completeness and, oftentimes, adequate validation; interactions among moderators is even more poorly understood. Attempts to improve significantly the empirical base for understanding moderator effects will be extremely costly and time-consuming. We have argued that the absence of realistic individual differences will compromise the validity of simulations for training and analysis. This argument is based on the generally held belief that the more realistic and representative the simulation, the more valid the training. While this is probably true in general, several questions remain.

- How much realism and representativeness is required to achieve desired levels of training or to achieve needed levels of analytic validity?
- Along what dimensions of human performance are realism and representativeness most critical?

UNCLASSIFIED

- How do these dimensions vary with the specific application?
- More generally, in what kinds of applications are realism and representativeness most critical?
- At what point does increased realism begin to detract from the training objectives.

Given the enormous costs associated with developing a firm empirical foundation of understanding the effects of moderators on performance, these questions become very important. We discuss this issue in more detail in the sections ahead.

Modeling the Behavior of Groups, Teams, and Organizations

Earlier we noted that the formal modeling of social structures, ranging from groups to teams and organizations, is in a very preliminary stage of development. The 2000 NATO report lists a number of factors and issues unique to social structures that need extensive investigation. They include the following.

- (1) Performance measurement. We currently lack good metrics of performance and behavior for military teams and organizations or for related non-military groups. This is a serious shortfall: the absence of quantitative measurement makes it difficult to assess the impact of training on military aggregates and, of equal importance, limits our ability to validate models.
- (2) Communication and co-ordination. Communication and co-ordination are essential features of effective teams. These processes have been studied extensively in laboratory and industrial settings, but much less in military settings. In particular, we do not have a good understanding of the impact of alternative organizational structures (e.g., “de-layering”), leadership styles or other moderator variables on communication and co-ordination in the military.
- (3) Co-operation, cohesion, and shared goals. These concepts relate to the degree to which members of a team or organization hold a shared understanding of the

UNCLASSIFIED

current situation and share common goals. The concepts have been studied to some extent, but primarily in the context of small, non-military teams. Our very limited ability to measure and model them is a significant shortfall given their importance to the effectiveness of military aggregates at all echelons and their sensitivity to both group and individual performance moderators.

(4) Group moderators. Two classes of moderators can exert an impact on the performance of a group, team, or organization. The first class influences the performance of individuals directly—for example fog, darkness or smoke can influence target acquisition—while the other class has an impact on the interactions among individuals—for example loss of shared goals through declining morale. Although the first class of moderators affects individual performance directly, it can affect the performance of the team through interactions among individuals. For example, if the target acquisition performance of individuals in a team is impaired by environmental factors, this is likely to affect the mission performance of the team as a whole. The second class of moderators operates directly on interactions among individuals and thus affects the behavior of the whole system. Though potentially of great significance, neither class of moderators has been studied systematically in the military context or in social structures beyond small groups.

Technologies Supporting Usability of Models and Simulations

Models of human behavior support many purposes and are used in diverse ways. Consequently, the process for developing these models is highly variable, limiting the generalizability and reusability of the models. In the paragraphs below, we will consider shortfalls of modeling practices, identifying the major sources of the problem and suggesting possible approaches to addressing them.

However, before we begin the discussion of models and their usability, let us focus a moment on human performance modeling within the context of a software development lifecycle. As Lawrence Peters observed, software engineering is at the

UNCLASSIFIED

intersection between theory and practice. This observation is highly applicable to the arena of human performance modeling and simulation.

Within the larger context of modeling and simulation, the introduction of human performance modeling is a relative newcomer. Typically, software development is described as a series of steps or phases—each with items to accomplish and a relative position with respect to other phases. For purposes of this report, the software development lifecycle begins with a Problem Description, followed by System Specification, Design, Coding and Testing. During the Problem Description phase the systems engineer gathers data to develop a fundamental understanding of the problem. The system specification phase focuses on the formalization of the data into a concise and clear statement of what the system is to do. The design phase results in a high level design showing how the system will accomplish its task; coding implements the refined design; and testing ensures that the system operates at some predetermined level of confidence.

Let's overlay some of the uniqueness of the human performance arena onto the software development cycle beginning with the first step, Problem Description. Currently, there is no standard process for analyzing the consequences of human performance on the problem, although some techniques in handbooks such as the Human Engineering Program Process and Procedures could be used to advantage. Consequently, human performance is not included in the fundamental understanding of what must be accomplished by the software. Task analysis techniques provide details about how work is performed; cognitive task analysis techniques yield the knowledge, thought processes and goal structure underlying the observed actions. Schraagen, Chipman, & Shalin (2000) recently reviewed the cognitive task analysis state-of-the-art and noted there is little in the CTA arena to assist the software developer in using the products of the task analysis to guide systems design. There have been previous attempts to store the output of task analysis efforts to enable reuse. One example is the CATT Database – a US Army funded activity focused on elements of Combined Arms Tactical Training. The relational database provided easy maneuverability across numerous field manuals and included

traceability among tasks, components, and definitions it has yet to achieve the anticipated degree of reuse.

Knowledge Acquisition (KA) is a key component of the second phase of the software development cycle, System Specification. KA is the process of obtaining knowledge from documentation or subject matter expert interviews (knowledge elicitation) with the goal of cataloguing that knowledge so it can be used in analyses. The current KA process is regarded as time consuming, costly, and brittle, being more of an art than a science. There is no underlying theory of KA, nor is there a generally agreed upon means for formatting and storing the information so that others may use it.

The elicitation of knowledge is often the acknowledged bottleneck in the process. Several reasons may be given for this, including:

1. Experts may not be aware of the knowledge they use in solving problems or the knowledge may be difficult to express in a useful manner. This makes it challenging for both for analysts to interview the experts and for experts to convey their knowledge to non-experts (in an attempt to satisfy the elicitor, the expert communicates whatever knowledge is accessible).
2. The methods used by analysts are time-consuming, largely relying on interviews, possibly giving a biased account of the domain expert's knowledge. This yields relatively unstructured information that is time-consuming to organize and does not support re-use of knowledge effectively. Furthermore, different methods may direct the expert's attention, producing different or inconsistent responses. Finally, individual analysts may have unique styles leading to differing results, even though using the same method or technique.
3. Knowledge is context-specific and task-specific, changing as tasks change. For instance, an expert's knowledge is, to a certain extent, bound to the existing systems and tasks on a particular platform. Re-use of this information for future platforms is risky if the systems differ significantly, implying that the

UNCLASSIFIED

KA process should be repeated for each new system under development.

Typically, the information is revised to accommodate the envisioned design, reducing the KA effort.

4. Integration of the KA information into software designs is currently difficult, so its usefulness is less than optimal. This may result in an iterative KA process—prototyping – further knowledge elicitation et cetera, without any clear sense of when the process is finished.

We will address each of the above-mentioned problems in terms of methods and techniques that are currently available to address these problems.

1. There are indirect techniques available for eliciting knowledge difficult to verbalize, for instance, pair-wise comparison scaling techniques. Although these techniques may elicit some underlying structure of the expert's knowledge, it may not be the structure used by the expert. These techniques yield limited types of knowledge they, the focus being on concepts and their relations rather than on heuristics, rules or strategies. For this reason, indirect techniques should be used in combination with process tracing techniques. In addition, the analyst requires some background knowledge in the study domain to use these techniques effectively. Experts may often regard these techniques with suspicion, due to their artificiality. Finally, the paired-comparison process is too cumbersome and time-consuming with any realistic number of concepts involved (more than 20). Therefore, these techniques have too many practical limitations to be generally useful for difficult knowledge elicitation.
2. Unstructured interviews are clearly not effective for eliciting knowledge. State-of-the-art empirical research would suggest relying more on well-crafted, realistic, scenarios for eliciting knowledge. The selection of cases presented to the expert is important: the cases should not be too routine (since the expert uses mainly skill-based knowledge that is difficult to quantify), nor too unfamiliar (since the expert can then only guess). KA is now generally viewed as a 'modeling' rather than a 'mining' activity; that is, it is accepted that the

UNCLASSIFIED

analyst comes up with a model of the expert's knowledge rather than claiming to have recovered the true nuggets of knowledge from the expert's head. Being a modeling activity implies that the result should be validated against some measure of effectiveness.

3. Re-use of knowledge is currently a difficult issue to resolve. Obviously, there are several layers of knowledge involved in any task, and some knowledge is more generalizable than is other knowledge. We need representations at different levels of abstraction in order to deal with these different kinds of knowledge.
4. There are a few promising approaches that may help use the KA results in the design process. One approach is to use generic task models, as developed in the KADS (Knowledge Acquisition, Documentation and Support) project by Wielinga and co-workers. This approach clearly separates the expert's knowledge from the representations used in the final implementation, thereby preventing a rapid prototyping approach, which is potentially costly. Second, system designers should evaluate the type of information required for a particular project first, and then select a CTA method that yields that type of information, (but see Allsopp, Harrison and Sheppard, 2001).

A major component of the testing phase in software development is validation. Establishing the validity of human behavior representations (HBRs) involves comparing a model or simulation result with corresponding human trial results to determine if the behaviors are sufficiently similar for a particular application. Without proper validation, no conclusion about the suitability of a model or simulation can be drawn with confidence. The responses of HBRs can be very sensitive to input variables, and even simple HBRs involve highly non-linear relationships between input variables and responses. This nonlinearity means a behavior observed for one set of conditions may not be produced for a similar set. The lack of well-established techniques and tools to support HBR validation makes these problems more difficult. Therefore, the HBR-community

needs to accept the challenge to develop both a sound method and a toolset to support the validation of human behavior models.

Current HBR validation practice relies primarily upon face validation by SMEs. During this process, an SME drives the HBR through the scenario space by issuing commands or changing the stimulating situation, and determines, often qualitatively, whether that behavior seems reasonable. Two resources exist to support HBR validation: (1) knowledge based system verification, validation, evaluation and testing (VVE&T) technology, and (2) human behavior science.

KBS VVE&T Technology

Significant investment has gone into the development of Knowledge Based Systems (KBS) to perform a variety of expert functions including diagnosis, decision support and automatic control. KBSs have been used in important applications such as flight control, financial management, disease diagnosis and treatment recommendation. These decision-support applications have driven the development of technology to measure and improve the validity of KBS behavior. HBRs are a class of KBS so the technology supporting the VVE&T of KBSs presents a tremendous and, heretofore, untapped resource for HBR validation. An extensive review of the KBS VVE&T literature shows it as a well-developed technology, with significant research and development activity over the last 27 years. It contains a rich body of theory, tools and techniques, described in numerous books, journal articles and conference and workshop proceedings that may apply to HBR validation under the right circumstances. This literature also revealed an enormous amount of experience in VVE&T of KBSs with 115 different references for diverse applications. By far, most of this experience related to medical applications where the results from any KBS can have life threatening consequences.

Of the KBS VVE&T technology references found, few directly addressed the evaluation of HBRs. The largest amount of work related explicitly to expert and decision support systems. Most of this work specifically addressed KBSs using production rule

UNCLASSIFIED

knowledge representations. While this focus upon rule sets may help validate some existing HBRs, none were designed to take advantage of these KBS VVE&T tools. As a result, changes to the HBRs or the tools will be necessary to benefit from these existing resources. Further, the VVE&T theory, techniques and tools apply only to the cognitive functions of HBRs and cannot be used for validation of the effects of behavior moderators or performance limitations. Nonetheless, KBS VVE&T techniques and tools may reduce the number of errors that exist in complex knowledge bases and integrated systems by automatically testing for completeness, consistency, coherence and redundancy.

Human testing of these knowledge bases is often tedious and subject to error. Automated KBS VVE&T tools could ease this task as well as locate areas where behavior is most complex, requiring SME intervention and concentrated results testing. This would improve the value of test results and the confidence in the entire representation.

Typically, SMEs validate HBR model performance through results testing but application of behavioral science techniques to HBR validation has significant limitations. Testing psychological and physiological correspondences requires extensive validated models of psychological and physiological phenomena. While many comprehensive psychological models exist, relatively few of them have been applied to HBR validation, especially for simulation applications. Like the physiological models, many psychological models deal with specific conditions limiting their usefulness for evaluating HBRs representing behavior for realistic situations. As psychological and physiological models become richer, their utility for HBR validation will increase. As with models and simulations of physical systems, model correspondence testing must be done at several levels of abstraction. Only consistent results between these different levels can guarantee validity; no single level of correspondence testing will likely be sufficient for any application.

During the transition from Description to Design and Coding in software development, the concept of composability becomes relevant. Composability is defined

as the ability to rapidly configure, initialize, and test an exercise by logically assembling a simulation from a pool of reusable elements. This concept is not unique to M&S – the software community has explored the use (and reuse) of components for many years. The objective of component-based simulation design is reduce the system construction effort by using existing components. This approach has some requirements. First, it needs a common approach for fully describing the components. Second, a model repository than can be easily used by analysts is needed to manage a broad range of components. Third, there must be a guide describing how to couple components and the implication for the resulting behaviour of their combined use. Finally, achieving this objective implies that metrics exist to validate the resulting simulation.

The critical first step in composability for human modeling is the creation of a generic framework, or body of understanding, for the design of repositories to archive models, model components, and algorithms. It must be possible to search the repositories efficiently, using multiple application-dependent variables such as level of resolution, computational requirements and interoperability as keywords.

Another concept that affects the Design and Coding phases of software development is interoperability. The capability to interconnect simulators and simulations has been successfully demonstrated in many different forums. The time seems right for scientists and engineers to support the HBR community by moving from stand-alone applications to networked applications, adopting standards such as the High Level Architecture. Integrated, interoperable models cannot be achieved merely by interconnecting component modules and considerable work remains to provide interoperability among HBR models. The keys to success lie in the community's ability to

- (1) identify the best means for assembling and sharing human behavior representations modules among simulation system developers
- (2) determine procedures for extracting components of human behavior representation modules that could be used in other modules

UNCLASSIFIED

- (3) determine procedures for passing data among modelers.

Development of an HLA-based testbed for HBR experimentation is also desirable to foster research into the issues affecting interoperability of human behavior representations.

Interoperability is multi-dimensional with many considerations

- Technical interoperability means that the implemented system must be easily integrated in the simulation environment or Command and Control system via protocols, communication links et cetera. Government and commercial developments like XML, CORBA, and RTI mainly cover this domain and should be considered for HBR.
- Interoperability on the data exchange level means that HBR and the environment must use the same representation schemes (data model et cetera). A standard approach is required that covers C2 and simulation systems, possibly based on different models developed in NATO (ATCCIS, NC3DM) for C2 systems. This is a field where research and development is required.
- Interoperability on the information exchange level covers not only requirements regarding data definition but also an understanding of the exchanged data (information). In this field, problems like aggregation and dispersion must be discussed. Another topic to be discussed is the relationship between HBR and simulation models including the re-use of HBR models in different simulations.

Roadmaps of Future Developments

Knowledge Acquisition Roadmap

Currently, knowledge acquisition leading to the development of human behavior models is considered to be mainly an art; knowledge acquisition efforts tend to be tied to specific applications and their underlying software architectures, thereby severely

UNCLASSIFIED

limiting their reuse in future applications. A number of developments are needed to enable reusability and to promote a more structured, scientific approach to knowledge acquisition, such as

- Development of initial data mining capabilities
- Development of databases with suitable indices, frames, and schemata attached
- Lessons learned by modellers in using structured databases
- Exploration of a scientific theory of human knowledge acquisition to promote knowledge acquisition analogues

Validation Roadmap

Research on validation of human behavior models in the short term (2002-2007) should identify current best practices and develop a systematic validation method. A number of efforts could help to address this point including

- Establish a process to develop testable requirements by using multi-disciplinary teams
- Develop an improved method for including SMEs in the validation process
- Define validation information requirements to be included with each HBR description
- Identify lessons learned from previous validation experiences
- Develop procedures for validation (e.g. from Recommended Practices Guide – DMSO)
- Identify examples of best practice

Composability and Interoperability Roadmap

In the short term (2002-2007), efforts should be made to develop

- A framework describing human behavior components to help determine integration

requirements

- A framework for knowledge representation that is independent of inference engine
- Interoperability on a semantic level, which is currently not given

References

Allsopp, D.J., Harrison, A., Sheppard, C. (2001) The representation and storage of military knowledge to support the automated production of computer generated force command agents. *Proceedings of the 10th Conference on Computer Generated Forces and Behavioral Representation*, pp. 159-169, Norfolk, VA

Dompke, U. & von Bayern, A (Eds) (2001) *Final report LTSS SAS-017 on Human Behaviour Representation* NATO R&T Organization, BP 25, 7 Rue ANCELLE, F-92201 NEUILLY-SUR-SEINE, CEDEX, France

Oser, R., McClusky, M.R., Blickensderfer, E., Campbell, G.E., & Lyons, D.M. (2000). Real-time modification of large scale exercises: Supporting the management of human trainer resources. *Proceedings of the 21st Interservice/Industry Training, Simulation and Education Conference*. pp.724-734.

Pew, R.W. & Mavor, A.S. (Eds) (1998) *Modeling Human and Organizational Behavior. Applications to military simulations*. Washington, DC: National Academy Press.

Schraagen, J.M, Chipman, S.F, & Shalin, V.L. (Eds) (2000) *Cognitive Task Analysis*. Mahwah NJ: Lawrence Erlbaum Associates.

Wald, B. (1998). ONR Simulation Study Team Final Report. Arlington, VA: Office of Naval Research.

Overall Conclusion and Recommendation

The most significant shortfalls in human performance modeling technology identified by AG-19 fall primarily within the training arena. As indicated in Chapter 4, these shortfalls lie in two broad technical areas: (1) the behavioral competence and the realism of human performance models, and (2) the technologies supporting their application. As a consequence of the first of these two shortfalls, models currently available for military simulations cannot provide warfighters with the worthy adversaries or able colleagues needed to achieve effective simulation-based training without the assistance of human controllers. Using human controllers to provide higher level cognitive capabilities lacking in simulated forces improves training, however, it is a very expensive solution that makes any-time, any-where, on-demand training virtually impossible. Consequently, a critical agenda item for the TTCP modeling and simulation community in the years ahead is to study techniques to enhance the behavioral and cognitive competence of human performance models along dimensions critical to military operations. The second shortfall in technologies supporting the application of human performance models is a significant cost of the military modeling and simulation enterprise. The techniques currently available for acquiring the knowledge needed to model the cognitive, perceptual and motor requirements of complex military tasks are extremely time-consuming, and need to be made more efficient. Moreover, we do not even have agreed-upon principles that allow us to determine what level of resolution is required to meet the requirements of new applications. Consequently, we are open to the problems both of over-resolution, which is unnecessarily expensive, or under-resolution, which promotes ineffective training. Added to this, we have not developed effective techniques to support the efficient re-use of models and model components across applications that differ in their specific focus and resolution requirements.

Since all the shortfalls we have identified in this report fall within the learning arena, we recommend that that TTCP follow-on activity be captured and managed under the TP-2 (Training) umbrella. It is not clear to the Action Group whether a separate Technical Panel would be able to effectively address the shortfalls we have identified.

UNCLASSIFIED

We recommend that a collaborative activity under TP-2 be established, starting with the participation at least of the US and CA, although it is hoped that UK, NZ and AU would also participate. The primary function of the collaborative activity we propose would be to allow relationships and discussion to continue among participating nations, focusing work on selected shortfalls that may lend themselves to international collaboration.

Annotated Bibliography of key recent reviews of the state-of-the-art in human performance and human behavior modelling

Modeling Human and Organizational Behavior: Application to military simulations

Richard W. Pew and Anne S. Mavor, editors

National Academy Press, 1998

This report is the work of the National Research Council's Panel on Modeling Human Behavior and Command Decision Making: Representations for Military Simulations. The panel was charged with reviewing the state of the art in human behavior representation as applied to military simulations, with emphasis on the challenging areas of cognitive, team, and organizational behavior. The project covered an 18-month period between 1996 and 1998.

Contents

Introduction

Human behavior representation: Military requirements and current models

Integrative architectures for modeling the individual combatant

Attention and multitasking

Memory and learning

Human decision making

Situation awareness

Planning

Behavior moderators

Modeling of behavior at the unit level

Information warfare: A structural perspective

Methodological issues and approaches

Conclusions and recommendations:

The panel's review has indicated that there are many areas today in which models of human behavior are needed and in which the models being used are inadequate. There are many other areas in which there is a clear need for models, but none exist.

UNCLASSIFIED

The overall recommendations of the panel are organized into two major areas—a programmatic framework and infrastructure/information exchange. The programmatic framework proceeds along four themes: collect and disseminate human performance data; create accreditation procedures for models of human behavior; support sustained model development in focused domains; and support theory development and basic research in relevant areas. The recommendations for infrastructure and information exchange are aimed at influencing and shaping modeling priorities within the services. These include collaboration, conferences, interservice communication, and education/training.

<http://www.nap.edu>

Modeling and Simulation: Linking entertainment and defense

Committee on Modeling and Simulation: Opportunities for Collaboration between the Defense and Entertainment Research Communities

Mike Zyda, Chair

National Academy Press, 1997

The National Research Council's Computer Science and Telecommunications Board convened a multidisciplinary committee to evaluate the extent to which the entertainment industry and the Department of Defense might be able to better leverage each other's capabilities in modeling and simulation technology and to identify potential areas for greater collaboration. This report represents the committee's attempt to capture key themes identified during a two-day workshop held October 19 and 20, 1996.

Contents

Introduction

Setting a common research agenda

Technologies for immersive simulated environments

UNCLASSIFIED

Networked simulation
Standards for interoperability
Computer-generated characters
Common research challenges
Adapatability
Modeling individual behaviors
Human representations
Aggregation and disaggregation
Spectator roles
Tools for creating simulated environments
Setting the process in motion

Concluding remarks:

Strong commonalities exist between defense and entertainment applications of modeling and simulation and in the technologies needed to support them. Whereas the Department of Defense has traditionally led the field and provided a significant portion of related funding, the entertainment industry has made rapid advances in 3D graphics generation, networked simulation, computer-generated characters, and immersive environments. Aligning the research agendas of these two communities offers the opportunity to more rapidly and economically achieve the goals of both.

<http://www.nap.edu>

Techniques for modelling human performance in synthetic environments: A supplementary review

UNCLASSIFIED

Frank E. Ritter, Nigel R. Shadbolt, David Elliman, Richard Young, Fernand Gobet and
Gordon D. Baxter

ESRC Centre for Research in Development, Instruction and Training, Technical Report
No, 62, University of Nottingham, Nottingham, UK, August 1999

In March 1999, DERA and DSTO sponsored a workshop that extended the NRC findings by (a) situating the results with respect to UK needs and capabilities and (b) exploring a few areas (e.g., usability) and architectures (e.g., Cogent, Jack) that were not previously available. The workshop participants examined current work related to modeling more complete human performance on cognitive modeling of emotion, advanced techniques for testing and building models of behavior, new cognitive architectures, and agent and Belief, Desires and Intentions technology. The report concludes with a list of high payoff projects for modeling human performance in synthetic environments.

Contents

Tasks and objectives for modeling behavior in synthetic environments

Summary of Pew and Mavor

Current objectives: More complete performance, more realistically

Current objectives: Improved usability

Summary of techniques for modeling behavior

Review of approaches with respect to objectives

Projects providing more complete performance

Projects supporting integration

Projects improving usability

UNCLASSIFIED

<http://www.psychology.nottingham.ac.uk/staff/ritter/mod-beh-sup-rev.pdf>

Final report LTSS/48 on Computer Generated Forces (CGF)

Uwe Dompke and G Scheckeler, editors

R&T Agency, NATO Headquarters, B-1110 Brussels, Belgium, July 1998

This final report is the culmination of discussions during a two-week Multinational Exercise on computer generated forces held in Alexandria, Virginia in June 1997. The NATO environment calls for increased cooperation between services as well as between nations, and for new concepts and systems. The objectives of LTSS/48 on CGF Technology were to identify technology areas that could support CGF based on military operational needs, and to give an assessment on the development of these technologies and their impact on CGF. The main conclusions from the study were: (1) computer generated forces is a set of technologies with broad Defense applications; (2) new applications of CGFs will need more intelligent and more autonomous behaviors; and (3) cost-effective development and use of CGFs depends on the ability to have shareable (among users, nations) representations and flexible system architecture. The report concludes with recommendations for increased multi-national program coordination and demonstration programs.

Contents

Introduction

Definition of CGF

Application Areas

 Training and exercises

 Defence planning

UNCLASSIFIED

Operations

Acquisition and procurement

Relevant technologies and military implications

Synthetic environment

Simulation system architecture

Modelling of CGFs

Human behaviour representation

Human/system interactions

Summary and conclusions

Major recommendations

Final report LTSS SAS-017 on Human Behaviour Representation

Uwe Dompke and Alex von Bayern, editors

R&T Organization, BP 25, 7 Rue ANCELLE, F-92201 NEUILLY-SUR-SEINE,
CEDEX, France

This final report is the culmination of discussions during a two-week Multinational Exercise on human behaviour representation held in Orlando, Florida in April 2000. The objectives of LTSS SAS-017 on HBR Technology were to identify technology areas based on military operational needs that could support HBR and to give an assessment on the development of these technologies and their impact on HBR. The relevant technologies discussed are individual behavior representation, team/group/organisational behahviour representation, validation, knowledge acquisition and interoperability/composability. The main conclusions from the study were: (1) HBR is a critical element for many military simulations; (2) currently available human behavior models yield over-simplistic behavior; (3) investment is needed to identify and prioritise required human modelling capabilities, to determine, organize and disseminate

UNCLASSIFIED

current and future behaviour databases, and to create a multinational test-bed for model refinement, interoperability assessment, and model validation, verification and accreditation. The report concludes with recommendations for increased multi-national program coordination and demonstration programs.

Contents

Introduction, background, scope and structure

Context and description of terms used

Military application

Current state of the art

Emerging developments

 Individual human behavior representation

 Team, group, organisation behaviour

 Model development process

Conclusions and recommendations

UNCLASSIFIED

List of Abbreviations

AAAV	Advanced Amphibious Assault Vehicle
AAII	Australian Artificial Intelligence Institute
AAW	Anti-Air Warfare
AAWC	Anti-Air Warfare Centre
ACC	Aerospace Combat Command (US)
ACO	Air Control Officer, Airspace Control Order
ACT-R	Adaptive Control of Thought-Rational http://act.psy.cmu.edu/
ACT-R/PM	Adaptive Control of Thought-Rational/Perceptual Motor
ADF	Australian Defence Force
ADMT-TD	Advanced Distributed Mission Trainer – Technology Demonstration
ADS	Air Defense Simulation
AEW&C	Airborne Early Warning and Control
AFAMS	Air Force Agency for Modeling and Simulation
AFC2TIG	Air Force Command and Control Training and Innovation Group
AG	Action Group (TTCP)
AIMP	Aurora Incremental Modernization Project
API	Application Programmer Interface
AOD	Air Operations Division (DSTO)
AMBR	Agent-based Modeling and Behavior Representation
ARI	Army Research Institute (US)
ASC	Army Simulation Centre

UNCLASSIFIED

ASW	Anti-Submarine Warfare
ATC	Air Traffic Control
ATO	Air Tasking Order
ATT	Aviation Team Training
AU	Australia
AWSIM	Air Warfare SIMulation (US)
BASS	Breakout Analysis Spread Sheet (US)
BDI	Belief Desire Intent
C2	Command and Control
C2DIF	Command and Control Data Interchange Format
C4ISR	Command, Control, Communication, Computers, Intelligence, Surveillance, Reconnaissance
CA	Canada
CAATS	Cognitive and Analytic Modeling Approaches for Tutoring Systems
CART	Combat Automation Requirements Testbed
CBS	Corps Battle Simulation
CCIS	Command and Control Information Systems
CF C2IS	Canadian Forces Command and Control Information System
CFEC	Canadian Forces Experimentation Centre (CA)
CCTT	Close Combat Tactical Trainer
CIS	Combat Instruction Set
CF	Canadian Forces
CF18	Canadian version of the A/F-18 A/B aircraft
CGF	Computer Generated Force

UNCLASSIFIED

CICO	Combat Information Center Officer
CHS	Centre for Human Sciences (formerly DERA, now QINETIQ)
COGNET	COGnition as a NEtwork of Tasks
CORBA	Common Object Request Broker Architecture
COTS	Commercial Off The Shelf
CPX	Command Post Exercise
C-PRS	C-Procedural Reasoning System
DARPA	Defense Advanced Research Projects Agency
DDD	Distributed Dynamic Decision making
DIS	Distributed Interactive Simulation
dMARS	Distributed Multi-Agent Reasoning System
DMSO	Defense Modeling and Simulation Office (US)
DMT	Distributed Mission Training
DND	Department of Defence (CA)
DoD	Department of Defense (US)
DRDC	Defence Research and Development Canada (CA)
D-OMAR	Distributed Operator Model Architecture
DSTO	Defence Science and Technology Organization (AU)
DV	Dependent Variable
EPIC	Executive Process Interactive Control
EWS	Electronic Warfare System
FEDEP	Federation Development and Execution Process (HLA)
FOM	Federation Object Model (HLA)
FSM	Finite State Machine

UNCLASSIFIED

GCI	Ground Controlled Intercept
GOMS	Goals, Operators, Methods and Selection rules,
GRADE	GRaphical Agent Development Environment
GTACS	Ground Theater Air Control System
GT-ASP	Georgia Tech Aegis Simulation Program
GUI	Graphical User Interface
HBR	Human Behavioral Representation
HEART	Human Engineering Analysis Requirements Tool
HFAT	Human Factors Analysis Tool
HFE	Human Factors Engineering
HIL,	Human-In-the-Loop
HLA	High Level Architecture, http://www.dmso.mil/
HPM	Human Performance Model
HQ	HeadQuarters
HUM	Human Resources and Performance (TTCP)
IA	Intelligent Agent
ICT	Institute for Creative Technologies, http://www.ict.usc.edu
IDP	Initial Development Phase
IFOR	Intelligent FORces
iGEN	Trademark name
I/ITSEC	Interservice/Industry Training, Simulation and Education Conference
IJCAI	International Joint Conference on Artificial Intelligence
IMCN	Intelligent Mission Controller Node
IP/PCT	Information Processing/Perceptual Control Theory

UNCLASSIFIED

IPME	Integrated Performance Modelling Environment, http://www.maad.com
JACK	JAVA Agent Compiler and Kernel
JAVA	Trademark name
JBASS	JSIMS Breakout Analysis Spread Sheet (US)
JFACC	Joint Forces Air Combat Commander
JORN	Jindalee Operational Radar Network
JSAF	Joint Semi-Automated Forces (US)
JSIMS	Joint SIMulation System (US)
JWARS	Joint Warfare System (US)
JWFC	Joint WarFighting Center (US)
KA	Knowledge Acquisition
KADS	Knowledge Acquisition, Documentation and Support
KBS	Knowledge Based System
NAWCTSD	Naval Air Warfare Center Training Systems Division
MAAD	Micro Analysis and Design Inc.
MALO	Maritime Air Littoral Operations
MAV	Micro-Air Vehicle
MDAPS	Major Defense Acquisition ProgramS
MHP	Maritime Helicopter Project
MoD	Ministry of Defence (UK)
ModSAF	Modular Semi-Automated Forces
MOUT	Military Operations in Urbanized Terrain
MNS	Mission Need Statement
MRE	Mission Rehearsal Exercise

UNCLASSIFIED

M&S	Modeling and Simulation
MTRS	Mission-Training and Rehearsal System
NASM	National Air and Space Model (US)
NAF	Numbered Air Force
NATO	North Atlantic Treaty Organization
NBC	Nuclear, Biologic, Chemical
NSI	Natural Selection Inc.
OA	Operations Analysis
OneSAF	One Semi-Automated Force
ONR	Office of Naval Research (US)
OR	Operations Research
ORD	Operational Requirements Document
PO	(ModSAF)
POP	Prediction of Operator Performance
R&D	Research and Development
RAAF	Royal Australian Air Force
RESA	Research, Evaluation, and Systems Analysis (US)
RO	Radar Officer
RPICN	Role Player Intelligent Controller Node
RPD	Recognition-Primed Decision Making
RTI	Run-Time Infrastructure (HLA)
SAF	Semi-Automated Forces
SAIC	Science Applications International Corporation
SAMPLE	Situation Awareness Model for Pilot-in-the-Loop Evaluation

UNCLASSIFIED

SAP	Situation Awareness Panel
SBA	Simulation Based Acquisition
SBIR	Small Business Innovate Research
SC	Soft Computing
SCOTT	Synthetic Cognition for Operational Team Training
SECO	Synthetic Environment Co-ordination Office (UK)
SEM	Sensor Employment Manager
SIMNET SAF	SIMulation NETwork Semi-Automated Force
SIMNET-T	SIMulation NETwork - Trainer
SME	Subject Matter Expert
SOAR	State, Operator And Result, http://ai.eecs.umich.edu/soar/
STEMM	Simulation of Team Mental Models
STOW	Synthetic Theatre of War
SuW	Surface Warfare
TADMUS	TAactical Decision-Making Under Stress (US)
TAMSS	Tactical Aviation Mission System Simulation (CA)
TACCO	Tactical Coordinator
TBMCS	Theater Battle Management Core System
T&E	Test and Evaluation
TCT	Time Critical Target
TP	Technical Panel (TTCP)
TTCP	The Technology Cooperation Program
UK	United Kingdom
US, USA	United States of America

UNCLASSIFIED

USAF	United States Air Force
USN	United States Navy
VAE	Virtual Air Environment
VVE&T	Verification, Validation, Evaluation and Testing
WARSIM	Warfighter's Simulation
WISSARD	What if Simulation System for Advanced Research and Development
WOC	Wing Operation Center
XML	Extensible Markup Language